



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE PSICOLOGÍA
DIVISIÓN DE ESTUDIOS PROFESIONALES

VELOCIDAD DE CAMBIO DEL ENTORNO
EN EL MODELO CLÁSICO DE REFORZAMIENTO

T E S I S

QUE PARA OBTENER EL GRADO DE:
LICENCIADO EN PSICOLOGÍA

PRESENTA:

VELÁZQUEZ VARGAS CARLOS ALAN

DIRECTOR DE TESIS:

DR. ARTURO BOUZAS RIAÑO

REVISOR:

DR. GERMÁN PALAFOX PALAFOX

SINODALES:

DR. GUSTAVO BACHÁ MÉNDEZ

DR. ÓSCAR ZAMORA ARÉVALO

DR. ÁNGEL EUGENIO TOVAR Y ROMO

PROYECTO PAPIME-PE310016



CIUDAD DE MÉXICO

ENERO, 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**Velocidad de cambio del entorno
en el modelo clásico de reforzamiento**

por

Velázquez Vargas Carlos Alan

Tesis presentada para obtener el grado de

Licenciado en Psicología

en la

FACULTAD DE PSICOLOGÍA

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad de México. Enero, 2018

A Arturo Velázquez

Agradecimientos

A la Universidad Nacional Autónoma de México, que me ha permitido formarme en sus aulas. A mis padres, por su confianza y apoyo constante. A mi hermano Arturo, por su curiosidad e intuición sobre el objetivo general de este trabajo. A mi director Arturo Bouzas, por su claridad y enseñanza a lo largo de estos años; por su rigor científico y su gran pasión, que me motivan a seguir mi vida académica. A Robert Wilson, que hizo posible que naciera este trabajo en la Universidad de Arizona y por su asesoría continua después de mi regreso a México. A Siyu Wang, por su gran amistad y apoyo en la programación y modelamiento en Matlab. A Manuel Villarreal, por las correcciones, aclaraciones y charlas que me enriquecieron enormemente. A Elena Villalobos, por su asesoría paciente, atención y motivación a lo largo del trabajo. A José Luis, por las pláticas y sugerencias entorno al modelamiento bayesiano. A mi revisor, Germán Palafox, y a mis sinodales, Gustavo Bachá, Ángel Tovar y Óscar Zamora por sus comentarios y correcciones. A mis compañeros del Laboratorio 25: Felisa, Uriel y Niño, y a los amigos que siempre me han acompañado: Daniel, Casandra, Alondra, Víctor y Arely.

Índice general

1. Introducción	1
2. Marco teórico	3
2.1. Predicción en entornos que varían momento a momento	3
2.2. El Filtro de Kalman como regla de aprendizaje	7
3. Método	9
3.1. Lugar y participantes	9
3.2. Materiales	9
3.3. Procedimiento	9
3.4. Descripción del experimento	10
3.5. Dinámica de las observaciones	11
3.6. Condiciones experimentales	11
4. Modelamiento jerárquico	13
5. Resultados	16
5.1. Estimación del proceso generador	16
5.2. Actualización ensayo a ensayo	18
6. Estimación paramétrica	20
7. Precisión del modelo	24

8. Discusión	28
8.1. Tasas de aprendizaje e incertidumbre interna asintóticas	29
8.2. Valores óptimos de los parámetros	29
8.3. Actualización ensayo a ensayo	31
8.4. Implicaciones para los modelos de ensayo y error	31
9. Conclusiones	33
10. Referencias	34

Velocidad de cambio del entorno en el modelo clásico de reforzamiento

por

Velázquez Vargas Carlos Alan

Resumen

El presente trabajo tiene como objetivo estudiar cómo las personas realizan predicciones en un entorno que fluctúa ensayo a ensayo en un marco de aprendizaje por reforzamiento. Se diseñó un experimento en computadora donde los participantes deben predecir la ubicación de un nave espacial que orbita alrededor del planeta Tierra. Su posición se genera a partir de una distribución normal con media que varía a lo largo de los ensayos y cuatro valores de varianza que definieron nuestras condiciones. Previo al análisis de datos, mostramos a través de simulaciones que la regla delta estándar en reforzamiento es incapaz de ajustarse adecuadamente en este tipo de entornos dada su estructura estática y, por lo tanto, una modificación del modelo es necesaria para describir las respuestas de los participantes. Con base en un algoritmo más general de reducción de error conocido como Filtro de Kalman, implementamos una regla delta que incluye un término de velocidad, el cual se actualiza vía error de predicción, para describir nuestros datos. En general, los resultados muestran que el modelo describe de manera precisa el comportamiento de los participantes en las condiciones experimentales. Este trabajo provee evidencia de que los algoritmos de aprendizaje por reforzamiento permiten describir comportamiento en entornos cambiantes. En particular, el modelo implementado sugiere que los participantes podrían estimar variables que controlan la dinámica del entorno, como la velocidad a la que cambia, y actualizar sus valores vía error de predicción.

Velocidad de cambio del entorno en el modelo clásico de reforzamiento

by

Velázquez Vargas Carlos Alan

Abstract

The current work aims to study how people make predictions in an environment that fluctuates from trial to trial under a reinforcement learning framework. A computer-based experiment was developed where subjects are required to predict the future location of a spaceship that orbits around planet Earth. Its position is sampled from a Gaussian distribution with mean changing over trials and four different values of variance that defined our noise conditions. Before analyzing participants' behavior, we showed via simulation that the standard delta-rule algorithm of reinforcement is unable to perform accurately in such a drifting task given its static nature, and therefore a modification of model is needed to describe responses of participants. Relying on a more general error-driven algorithm known as the Kalman Filter, we implemented a delta-rule model that includes a velocity term which is updated via prediction errors to fit our data. In general, our results provided support for the model as an accurate descriptor of participants' behavior in the experimental conditions. This work offers further evidence for reinforcement learning algorithms as plausible mechanisms of learning in changing environments. In particular, the implemented model suggests that people could keep estimates of variables that control the dynamics of the environment such as the velocity of change and to update their values via prediction errors.

Capítulo 1

Introducción

De particular interés en la literatura de aprendizaje es cómo las personas y otros animales actualizan sus creencias en entornos cambiantes. Tradicionalmente, los algoritmos de ensayo y error (o de reforzamiento) han permitido representar dicho proceso de actualización bajo condiciones estables. Sin embargo, en su forma más simple son incapaces de describir el comportamiento adecuadamente cuando las condiciones varían. En particular, si los cambios ocurren momento a momento, mostraremos vía simulación que el modelo clásico de reforzamiento (Rescorla y Wagner, 1972) no está diseñado para realizar predicciones correctas, por lo cual representa un modelo inapropiado para describir el comportamiento en tales circunstancias. Así mismo, proponemos que cuando las personas se enfrentan a un ambiente que se modifica momento a momento, y del cual solo tienen observaciones ruidosas, podrían utilizar una versión simplificada de un Filtro de Kalman (Kalman 1960) como regla de aprendizaje, la cual incorpora la estimación de la velocidad de cambio del entorno.

Con el fin de indagar si las personas implementan una regla de este tipo, desarrollamos una tarea en computadora donde nuestros participantes deben predecir la ubicación de una nave espacial que orbita alrededor de la tierra. Su posición se modifica dependiendo de un término de velocidad cuyo valor cambia a lo largo de los ensayos, y por ruido aleatorio (al que denominamos *incertidumbre externa*), dependiente de la condición experimental. Para describir el comportamiento de nuestros sujetos, implementamos una versión jerárquica de la regla de aprendizaje, la cual asume que los parámetros individuales provienen de distribuciones definidas por condición, y por lo tanto que son dependientes de la incertidumbre externa. Utilizamos

metodología de inferencia bayesiana para estimar los valores de los parámetros a nivel grupal e individual. Dicha aproximación nos permitió representar el nivel de credibilidad sobre cada valor paramétrico, así como un rango de predicciones del modelo. De manera general, nuestros datos sugieren que los participantes son capaces de predecir la posición más probable de la nave en la mayor parte de los ensayos, y que el modelo implementado puede capturar adecuadamente este comportamiento. La inferencia paramétrica sugiere que la posición y velocidad de la nave se actualizan ensayo a ensayo, y que a medida que la incertidumbre externa disminuye, los participantes reducen la incertidumbre de sus propias estimaciones, y confían más en cada nueva observación, lo cual se ve reflejado en el incremento de las tasas de aprendizaje. Lo anterior brinda evidencia a favor de los algoritmos de ensayo y error como reglas de aprendizaje plausibles que permiten a las personas actualizar sus creencias tanto en entornos estables como dinámicos.

Capítulo 2

Marco teórico

2.1. Predicción en entornos que varían momento a momento

Las decisiones con frecuencia tienen lugar en entornos que varían en el tiempo. La disponibilidad de alimento para algunos animales puede cambiar gradualmente en función del crecimiento de la fuente o del consumo constante; así mismo, la posición de los objetos en el espacio puede modificarse dada una tasa de desplazamiento determinada. El mantener estimaciones adecuadas en estas circunstancias permitiría a los organismos distribuir mejor su comportamiento, e.g. al cambiar a una fuente más rica de alimento cuando la tendencia de la actual es decreciente, o al predecir adecuadamente la posición de un objeto que se dirige peligrosamente hacia ellos. Para realizar esta tarea deben contar con una representación útil del entorno que les permita guiar decisiones futuras. En el caso más sencillo, esta representación es un número relacionado con la variable biológicamente importante (e.g. posición o disponibilidad de comida) que se actualiza conforme se obtiene más experiencia para reducir el error de predicción¹. Dicho proceso constituye la idea central del modelo clásico de reforzamiento, el cual ha sido utilizado ampliamente para describir aprendizaje en organismos (Para una revisión ver Bouzas, 2017) y en máquinas (Sutton y Barto, 1998) vía ensayo y error. Una forma común de representarlo es como una regla Delta:

¹La discrepancia entre lo que se predice y lo que en realidad se observa.

$$\begin{aligned}\delta_t &= z_t - \mu_t \\ \mu_{t+1} &= \mu_t + \alpha\delta_t\end{aligned}\tag{2-1}$$

donde, en un ensayo t , el error de predicción δ_t , ponderado por la tasa de aprendizaje α , es utilizado para actualizar la estimación presente μ_t , una vez que el resultado z_t es observado. Evidencia conductual (Bush y Mosteller, 1951; Rescorla y Wagner, 1972; Miller et al., 1995) y neurofisiológica (Schultz, 1997; Daw, 2013) sugiere que la regla Delta es un mecanismo de aprendizaje biológicamente plausible en mamíferos. Sin embargo, una de sus limitaciones es la incapacidad de describir comportamiento en entornos cambiantes. Cuando el entorno se modifica abruptamente, tanto animales (Baum, 2010; Gallistel et al., 2001) como humanos (Gallistel, 2014) son capaces de detectar el cambio rápidamente y ajustar su comportamiento a las nuevas demandas. Sin embargo, dada la naturaleza fija de α , la regla Delta es incapaz de capturar este patrón. Una α pequeña hace que las predicciones durante periodos de estabilidad sean correctas, pero provoca una adaptación lenta después de un cambio. Una α elevada tiene el efecto contrario; ajustándose ineficazmente durante períodos de estabilidad, pero respondiendo rápidamente en presencia de cambios. Ante esta dificultad, algunas de las soluciones reportadas en la literatura incluyen utilizar una tasa de aprendizaje adaptativa (Nassar, 2010) o múltiples reglas delta con su propia tasa de aprendizaje (Wilson, 2013), permitiendo un ajuste eficaz tanto a los períodos de estabilidad como ante la presencia de un cambio.

Cuando el entorno en lugar de modificarse abruptamente lo hace momento a momento y de manera gradual, animales (Innis y Staddon, 1971) y humanos (Ricci y Gallistel, 2017; Speekenbrink y Shanks, 2010) también son capaces de responder de forma adecuada siguiendo la dirección de las variaciones. Sin embargo, aún no es claro si dicho proceso podría también computarse vía ensayo y error. Ricci y Gallistel (2017) muestran que al estimar el parámetro de una distribución Bernoulli cuyo valor se modifica ensayo a ensayo, las personas varían sus predicciones siguiendo un patrón escalonado, análogo al ajuste a cambios abruptos, y sugieren que el proceso de actualización se llevaría a cabo únicamente cuando existe suficiente evidencia de que las observaciones son distintas de las estimaciones, y no ensayo a ensayo como en una regla Delta. Por otro lado, Speekenbrink y Shanks (2010) muestran que cuando el peso que tiene un estímulo sobre las consecuencias varía gradualmente a lo largo del experimento, las personas

actualizan sus estimaciones de dichos valores de manera continua, análogo a un proceso de ensayo y error.

Si las personas y otros animales implementan reglas de ensayo y error o no para realizar estimaciones de su entorno cuando éste cambia gradualmente aún es motivo de debate (Gallistel, 2012; Ricci y Gallistel, 2017). Sin embargo, al tomar una regla Delta simple como modelo de este proceso, notamos que sus predicciones resultan inadecuadas. Consideremos el ejemplo mencionado al inicio donde la disponibilidad de alimento varía en el tiempo. La Figura 2-1 muestra una simulación donde, en un día determinado, cada valor (puntos azules) proviene de una distribución Normal con varianza 1 y media x (línea negra punteada) que varía de acuerdo $x_{t+1} = x_t + v_t$ donde v_t es un término de velocidad que sigue una carrera aleatoria de tiempo discreto ². En el panel superior se muestran las predicciones de una regla Delta (línea roja) utilizando la secuencia simulada de valores y $\alpha = 0.2$. En este caso, la adaptación del algoritmo es lenta dado que las estimaciones del pasado reciben un mayor peso en las nuevas predicciones, lo cual provoca que con el paso de los ensayos el algoritmo se vaya quedando atrás. En el panel inferior se muestran las predicciones de una regla Delta utilizando las mismas observaciones pero con $\alpha = 0.9$. Aunque el desempeño es mejor que en el caso anterior, dado el valor elevado de la tasa de aprendizaje las nuevas predicciones son similares al último dato observado, es decir, al valor del día anterior. Como resultado, cuando el proceso generador disminuye sus valores (día 12 al 17) o los incrementa (día 18 a 22) el modelo se adapta como si estuviera siempre un día atrás. De manera general, una regla Delta simple es incapaz de predecir la dirección y la magnitud a la que varían las observaciones independientemente del valor de α . En el mejor de los casos, utilizar una tasa de aprendizaje elevada genera predicciones que se encuentran un momento atrás.

Una forma de resolver este problema es contar con un estimado de la tendencia que siguen las observaciones. Por ejemplo, que tan rápido están cambiando y en qué dirección. Esta cantidad puede actualizarse conforme se obtiene más evidencia siguiendo la estructura de una regla Delta, e integrarse de manera lineal a la predicción final. Un modelo de esta naturaleza nos lleva a una versión simple del Filtro de Kalman (Kalman, 1960; Sutton, 1992). Estudios previos han

²Este proceso permite generar una secuencia en la que cada valor es igual al anterior más ruido aleatorio. Cuando los valores se generan en particiones arbitrarias del tiempo como son los ensayos, se denomina carrera aleatoria de tiempo discreto.

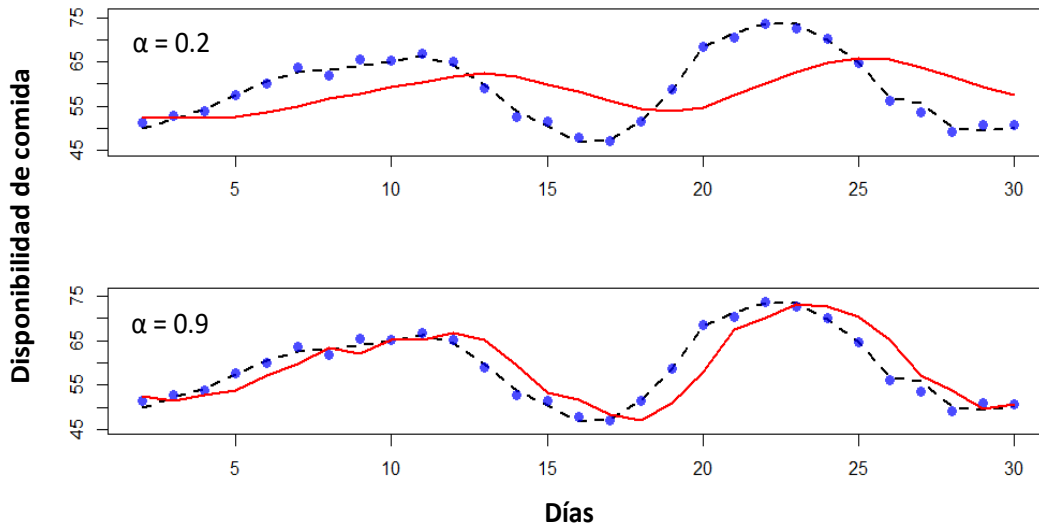


Figura 2-1: Predicciones de la regla Delta estándar (línea roja) con base en observaciones ruidosas (puntos azules) de la disponibilidad de comida (línea negra) cuando ésta se modifica a lo largo de los días. Superior: regla Delta con $\alpha = 0.2$. Inferior: regla Delta con $\alpha = 0.9$.

mostrado que el Filtro de Kalman es una herramienta útil para modelar procesamiento sensorial (Rao y Ballard, 1997, 1999), así como la adquisición y extinción de respuestas (Kakade y Dayan, 2000, 2002). En particular, podemos decir que constituye un algoritmo de ensayo y error, pero en múltiples dimensiones. En este trabajo nos preguntamos si una regla de aprendizaje similar al Filtro de Kalman puede predecir los patrones de cambio conductual cuando el entorno se modifica gradualmente y solo se tienen observaciones inciertas. Dicho modelo asume que los sujetos: a) estiman la velocidad de cambio de su entorno y la integran a sus predicciones, b) modulan el impacto de las nuevas observaciones vía tasas de aprendizaje dependiendo del grado de incertidumbre en ellas (incertidumbre externa), y c) acompañan cada estimación con cierto grado de error (incertidumbre interna), el cual depende a su vez del nivel de incertidumbre externa.

Para responder esta pregunta, desarrollamos una tarea en computadora donde nuestros participantes deben estimar la posición de una nave espacial que orbita alrededor de la tierra en función de un término de velocidad y ruido aleatorio. Observamos el grado en que nuestro modelo es capaz de describir las respuestas de los participantes, así como el comportamiento de los parámetros a nivel individual como grupal para evaluar si los cálculos propuestos por dicho

modelo son un posible mecanismo de aprendizaje en este tipo de entornos.

2.2. El Filtro de Kalman como regla de aprendizaje

Al resolver el problema de predicción, asumimos que los participantes integran en un valor la posición que tendrá la nave espacial en el momento siguiente con base en observaciones previas. El método óptimo para realizar esta tarea cuando las observaciones están contaminadas con ruido Normal, es el Filtro de Kalman (Kalman, 1960). Además de una herramienta matemática y estadística, es útil como modelo de aprendizaje (Grewal, 2014; Speekenbrik, 2010; Daw, 2009; Geist, 2008; Gershman, 2015, 2017). De manera general, consta de dos sencillos pasos. El primero, es generar estimaciones de las posibles variables que describen el estado del sistema. En particular, el sistema del experimento se refiere a la nave espacial, y el estado, a las variables que la describen en el espacio, i.e. en la pantalla. Por ejemplo, su posición, su velocidad o su aceleración en un momento determinado. Posteriormente, se actualizan dichos valores al adquirir nuevas observaciones. En este punto, se pondera la nueva información dependiendo del nivel de incertidumbre existente. Como se describe en el siguiente capítulo, el estado que define a la nave en todo momento consta de dos variables (su posición y su velocidad) y, por lo tanto, asumiremos que nuestro modelo mantiene estimaciones únicamente de ellas.

Suponga que se tiene los valores iniciales \hat{M}_t y \hat{V}_t para la posición y velocidad de la nave, respectivamente, en el tiempo t . Posteriormente, se obtiene una observación z_t , es decir, aparece la nave en la pantalla. El proceso de actualización está dado por:

$$M_t = \hat{M}_t + \alpha(z_t - \hat{M}_t) \quad (2-2)$$

$$V_t = \hat{V}_t + \beta(z_t - \hat{M}_t) \quad (2-3)$$

donde M_t y V_t son los valores independientes de la posición y la velocidad después de la observación z_t . Es importante enfatizar que M_t representa el valor de la posición *sin tomar en cuenta* el cambio por la velocidad. α y β representan las tasas de aprendizaje de cada variable, y conforman la denominada Ganancia de Kalman. Si su valor es cercano a 0, indica que la incertidumbre externa es alta y, por lo tanto, se les asignará poca importancia a las observaciones en el proceso de actualización. Como resultado, los estimados del algoritmo se mantienen estables. Si su valor es cercano a 1, indica que la incertidumbre externa es baja, y

por lo tanto se les dará un mayor peso a las observaciones en la actualización. En este caso, las estimaciones se vuelven inestables.

El siguiente paso es integrar los valores de las ecuaciones 2-2 y 2-3 para generar nuevas predicciones:

$$\hat{M}_{t+1} = M_t + \Delta t V_t \quad (2-4)$$

$$\hat{V}_{t+1} = V_t \quad (2-5)$$

donde \hat{M}_{t+1} y \hat{V}_{t+1} representan las predicciones del modelo para posición y velocidad, respectivamente, en el momento siguiente. Δt indica cada cuando el algoritmo actualiza sus predicciones. En este trabajo asumimos que lo hace cada ensayo y por lo tanto es igual a uno. Notar que \hat{M}_{t+1} considera el cambio debido a la velocidad, a diferencia de M_t en la ecuación 2-2.

Expresando las ecuaciones 2-4 y 2-5 en términos de 2-2 y 2-3 obtenemos:

$$\hat{M}_{t+1} = \hat{M}_t + \hat{V}_{t+1} + \alpha(z_t - \hat{M}_t) \quad (2-6)$$

$$\hat{V}_{t+1} = \hat{V}_t + \beta(z_t - \hat{M}_t) \quad (2-7)$$

De la ecuación 2-6 se tiene que la posición estimada de la nave en el momento siguiente está definida por la estimación previa, más el valor de la velocidad, más el error de predicción de la última estimación ponderado por la tasa de aprendizaje. La ecuación 2-7 está definida de forma similar, sin embargo, es independiente de las estimaciones de la posición, además cuenta con su propia tasa de aprendizaje. Se puede apreciar que la ecuación 2-6 es simplemente una regla delta que incluye el término \hat{V}_{t+1} . Si $\hat{V}_{t+1} = 0$, obtenemos el modelo clásico de reforzamiento (Bush y Mosteller, 1951).

Por último, asumimos que cada nueva estimación es acompañada de cierto error, lo cual nos permite representar al comportamiento como:

$$B_{t+1} \sim Normal(\hat{M}_{t+1}, \sigma^2) \quad (2-8)$$

donde B representa la respuesta del participante en el momento $t + 1$, y σ^2 el parámetro de error que denominaremos *incertidumbre interna*, el cual corresponde con la varianza de la distribución Normal.

Capítulo 3

Método

3.1. Lugar y participantes

El experimento se llevó acabo en la Unidad para el Desarrollo de Material de Enseñanza y Apropriación Tecnológica (UDEMAT) de la Facultad de Psicología de la Universidad Nacional Autónoma de México. Se recolectaron datos de un total de 72 individuos (55 mujeres) con una edad media de 19.83 años ($DE=2.035$).

3.2. Materiales

Se programó la tarea utilizando el software comercial Matlab® 2016b (MathWorks, 2016) y la extensión Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007), para el desarrollo de estímulos visuales. Se utilizó una pantalla de dimensiones 1920 x 1080 pixeles, y un ratón y teclado estándar de computador para navegar en el experimento. El sistema operativo utilizado fue Windows 10 de 64-bit.

3.3. Procedimiento

Los participantes fueron contactados vía correo electrónico, medio por el cual se les informó que podían ganar un total de \$400, \$200 y \$100 pesos mexicanos, si obtenían el primero, segundo o tercer mejor desempeño en la tarea, respectivamente. Posteriormente, se acordó un horario para asistir a la sesión experimental. Una vez presentes en el lugar, se proporcionó la hoja de

consentimiento informado a cada participante, y posterior a su aceptación se inició la tarea. Cada sesión estuvo integrada por un total de 5 a 8 participantes asignados a una computadora cada uno, y tuvo una duración promedio de 45 minutos.

3.4. Descripción del experimento

Se diseñó una tarea interactiva en computadora, en la cual se presenta la imagen de la tierra en el espacio, y una nave orbitando alrededor de ella. La Figura 3-1 es un esquema de la tarea.

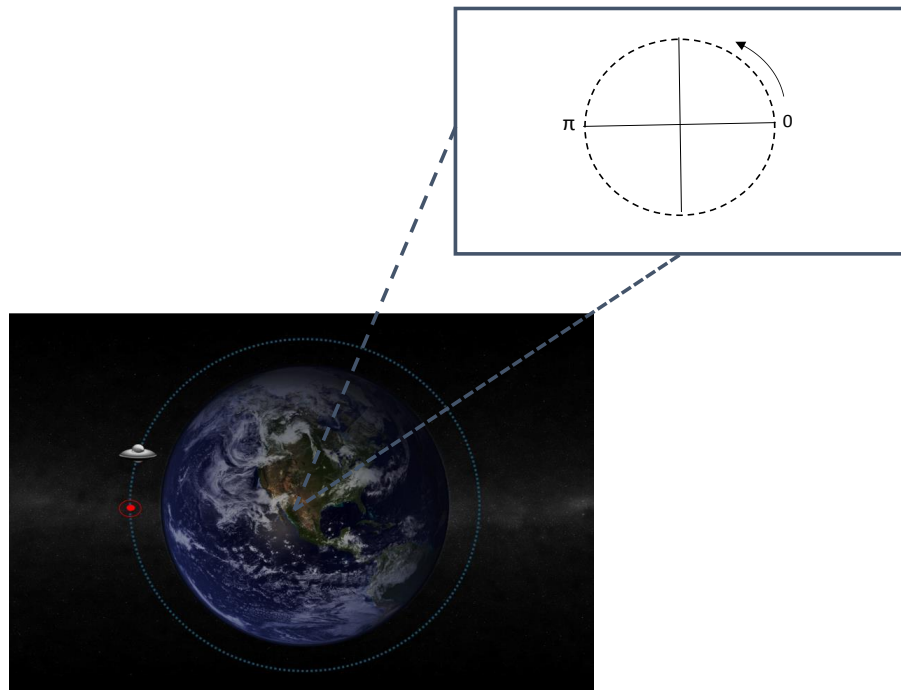


Figura 3-1: Diseño de la tarea

En cada ensayo, la nave se muestra por un tiempo breve en algún punto de su trayectoria (indicada con una línea azul punteada) y después desaparece. El participante debe entonces indicar con su cursor el lugar donde creé que se moverá en el siguiente ensayo. Si acierta a su nueva ubicación, la nave se torna de color rojo, si no acierta, permanece inalterable. Un punto rojo indica el lugar seleccionado en la trayectoria y un círculo punteado del mismo color indica el margen de error permitido (i.e. si la nave reaparece dentro de él, la predicción es considerada

exitosa). Esta secuencia se repite por un total de cuatro condiciones de trescientos ensayos cada una. Al finalizar cada condición, se muestra un mensaje al participante de que puede tomar un breve descanso y continuar cuando esté listo. En la parte superior derecha de la Figura 3-1 se observa la estructura de la tarea. Es importante señalar que los participantes no tenían acceso a dicha representación. La posición de la nave está dada en radianes (rad), utilizando el centro de la órbita como referencia, y un punto de partida marcado con un cero. Si la nave completa una vuelta sobre su órbita en dirección opuesta a las manecillas del reloj, es decir, al moverse un total de 2π rad, la segunda vuelta comenzaría con valores de 2π y hasta 4π rad, la tercera de 4π hasta 6π rad, y así sucesivamente. De manera similar, si la nave cruza el punto de partida en dirección a las manecillas del reloj, su posición estaría dada de acuerdo a los valores de la vuelta anterior. Seguimos la misma lógica para registrar las respuestas de los participantes. Esto permitió generar un rango de posibles posiciones que abarcaba de $-\infty$ a ∞ rad.

3.5. Dinámica de las observaciones

La tarea de los participantes consistió en estimar la posición verdadera de la nave x en el ensayo t , con base en observaciones ruidosas $z_t = x_t + \xi_{zt}$, donde ξ_{zt} es ruido aleatorio proveniente de una distribución normal con media en cero y varianza σ_z^2 y el modelo de su dinámica:

$$v_t = v_{t-1} + \zeta_{vt} \quad (3-1)$$

$$x_{t+1} = x_t + v_t \quad (3-2)$$

donde v_t es un término de velocidad y ζ_{vt} es ruido aleatorio proveniente de una distribución normal con media en cero y varianza σ_v^2 .

3.6. Condiciones experimentales

Los participantes llevaron a cabo la tarea en cuatro condiciones experimentales que diferían en la razón señal-ruido $\frac{\sigma_v^2}{\sigma_z^2}$. Intuitivamente, este valor nos indica el grado en que es posible discriminar el proceso generador. Si $\frac{\sigma_v^2}{\sigma_z^2} < 1$, existe un mayor nivel de ruido que de señal, y por lo tanto es más difícil de discriminar; si $\frac{\sigma_v^2}{\sigma_z^2} > 1$, la cantidad de señal es mayor que la de ruido, y por

lo tanto, es más sencillo de discriminar. Esta razón, es a su vez un indicador de la incertidumbre externa o grado de aleatoriedad del entorno. Posterior a una prueba piloto, se fijó un valor único de σ_v^2 para las cuatro condiciones de tal manera que la velocidad de la nave no incrementara al grado de parecer que se movía aleatoriamente. Con σ_v^2 fijo, el valor de σ_z^2 es entonces aquel que satisface la razón. El orden de las condiciones fue aleatorio para todos los participantes. La Tabla 3-1 muestra los valores utilizados por condición.

Tabla 3-1: Diseño del experimento

Condición	$\frac{\sigma_v^2}{\sigma_z^2}$	σ_v^2	σ_z^2
1	0.05	0.0049	0.098
2	0.5	0.0049	0.0098
3	1	0.0049	0.0049
4	2	0.0049	0.00245

Nota. Las unidades se encuentran en radianes.

Capítulo 4

Modelamiento jerárquico

Desarrollamos un modelo jerárquico con base en la regla de aprendizaje descrita en el capítulo 2 para analizar el comportamiento de los participantes. Los modelos jerárquicos asumen que los parámetros individuales son a su vez producto de un proceso de orden superior (Shiffrin et. al, 2008; Lee, 2016). Una forma simple de representarlo, es plantear que provienen de distribuciones Normales. En nuestro modelo, dichas distribuciones están definidas a nivel de las condiciones, sugiriendo que los participantes utilizan valores diferentes de los parámetros dependiendo del nivel de incertidumbre externa. Formalmente podemos decir que:

$$\hat{\alpha}_{kj} \sim Normal(\mu_k^{\hat{\alpha}}, \sigma_k^{2\hat{\alpha}}) \quad (4-1)$$

$$\hat{\beta}_{kj} \sim Normal(\mu_k^{\hat{\beta}}, \sigma_k^{2\hat{\beta}}) \quad (4-2)$$

$$\sigma_{kj}^2 \sim Normal(\mu_k^{\sigma^2}, \sigma_k^{2\sigma^2}) \quad (4-3)$$

donde en la condición k para el participante j , $\hat{\alpha}_{kj}$ y $\hat{\beta}_{kj}$ son aproximaciones de las tasas de aprendizaje para posición y velocidad, respectivamente, y σ_{kj}^2 la incertidumbre interna. En particular, $\hat{\alpha}_{kj}$ y $\hat{\beta}_{kj}$, se encuentran en el intervalo (0,1) como está especificado teóricamente en los modelos de ensayo y error. Para representar esta restricción, cada valor de $\hat{\alpha}_{kj}$ y $\hat{\beta}_{kj}$, es ajustado por la función de densidad acumulada de una distribución Normal estándar expresada por ϕ . Entonces, tenemos que:

$$\alpha_{kj} = \phi(\hat{\alpha}_{kj}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\hat{\alpha}_{kj}} e^{-\frac{\hat{\alpha}_{kj}^2}{2}} \quad (4-4)$$

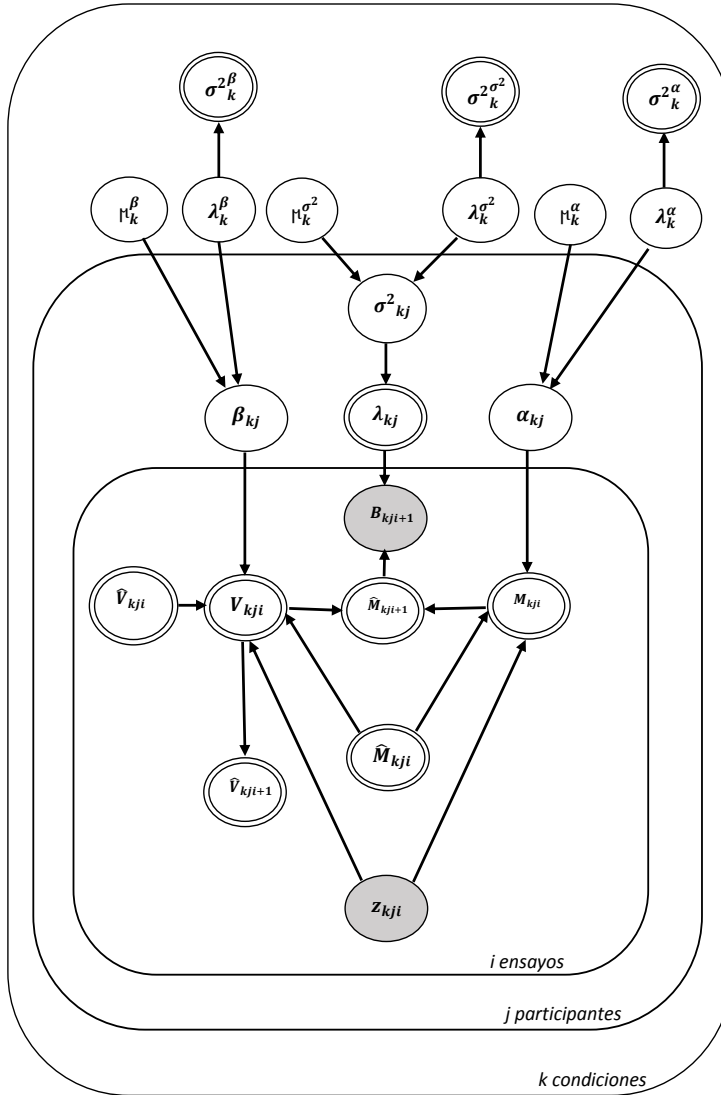
$$\beta_{kj} = \phi(\hat{\beta}_{kj}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\hat{\beta}_{kj}} e^{-\frac{\hat{\beta}_{kj}^2}{2}} \quad (4-5)$$

donde α_{kj} y β_{kj} representan los valores de las tasas de aprendizaje para posición y velocidad definidos para el intervalo (0,1), en la condición k para el participante j . σ_{kj}^2 está definida en el intervalo (0,∞) . Sin embargo, previo a la estimación de sus valores para cada participante utilizando estadística bayesiana (Capítulo 6), realizamos la estimación vía máxima verosimilitud, la cual requiere una menor cantidad de tiempo y permite tener una noción de los posibles valores de los parámetros. Los resultados indicaron que los valores máximos de σ_{kj}^2 se encontraban alrededor de 0.5. De esta manera, decidimos que un rango razonable para el parámetro sería (0,1) y, por lo tanto, que podríamos utilizar la función ϕ para restringirlo al igual que con las tasas de aprendizaje. Por lo tanto, σ_{kj}^2 se obtiene mediante:

$$\sigma_{kj}^2 = \phi(\hat{\sigma}_{kj}^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\hat{\sigma}_{kj}^2} e^{-\frac{\sigma_{kj}^4}{2}} \quad (4-6)$$

μ_k^α , μ_k^β y $\mu_k^{\sigma^2}$ representan medias de las distribuciones a nivel de condición, y $\sigma_k^{2\alpha}$, $\sigma_k^{2\beta}$ and $\sigma_k^{2\sigma^2}$ sus varianzas. Tanto medias como varianzas a nivel de condición reciben el nombre de hiperparámetros. En la Figura 4-1 se muestra una representación gráfica del modelo siguiendo la notación presentada en Lee y Wagenmakers (2013). Las variables se encuentran definidas dentro de nodos y las flechas indican relaciones de dependencia. Si el nodo es gris, indica que la variable es conocida, si es blanco, que es desconocida; si es un círculo, indica que la variable es continua, y si es un doble círculo, que es continua y está determinada. Los platos indican replicaciones independientes de un proceso dentro del modelo. A un costado del gráfico se detalla la dependencia de las variables y la distribución conocida de los parámetros antes de observar los datos.

Dada la condición k para el participante j , el proceso de inferencia del modelo en la figura 4-1 comienza con la i -ésima observación de la nave espacial, z_{kji} , que alimenta las reglas Delta M_{kji} y V_{kji} para posición y velocidad, respectivamente. Cada regla Delta cuenta con su propia tasa de aprendizaje: α_{kj} correspondiente a M_{kji} y β_{kj} a V_{kji} . Ambas están definidas para cada participante j y provienen de distribuciones normales definidas en la condición k . Posteriormente, se genera una nueva predicción de la posición \hat{M}_{kji+1} , sumando los valores de M_{kji} y V_{kji} , y una nueva predicción para velocidad \hat{V}_{kji+1} que es idéntica a V_{kji} . Finalmente, el comportamiento observado, B_{kji+1} , se genera de una distribución normal con media \hat{M}_{kji+1} y varianza σ_{kj}^2 ; ésta última definida para cada participante j , pero que se asume proviene de una distribución normal



$$\begin{aligned}
\mu_k^\alpha, \mu_k^\beta, \mu_k^{\sigma^2} &\sim \text{Normal}(0, 0.001) \\
\lambda_k^\alpha, \lambda_k^\beta, \lambda_k^{\sigma^2} &\sim \text{Gamma}(0.001, 0.001) \\
\sigma_k^{\sigma^2} &\leftarrow 1/\lambda_k^{\sigma^2} \\
\sigma_k^{2\beta} &\leftarrow 1/\lambda_k^\beta \\
\sigma_k^{2\sigma^2} &\leftarrow 1/\lambda_k^{\sigma^2} \\
\hat{\alpha}_{kj} &\sim \text{Normal}(\mu_k^\alpha, \lambda_k^\alpha) \\
\hat{\beta}_{kj} &\sim \text{Normal}(\mu_k^\beta, \lambda_k^\beta) \\
\hat{\sigma}_{kj}^2 &\sim \text{Normal}(\mu_k^{\sigma^2}, \lambda_k^{\sigma^2}) \\
\alpha_{kj} &\leftarrow \Phi(\hat{\alpha}_{kj}) \\
\beta_{kj} &\leftarrow \Phi(\hat{\beta}_{kj}) \\
\sigma_{kj}^2 &\leftarrow \Phi(\hat{\sigma}_{kj}^2) \\
\lambda_{kj} &\leftarrow 1/\sigma_{kj}^2 \\
V_{kji} &= \begin{cases} \hat{V}_{kji} + \beta_{kj}(z_{kji} - \hat{M}_{kji}) & \text{if } i > 1 \\ 0 & \text{if } i = 1 \end{cases} \\
M_{kji} &= \begin{cases} \hat{M}_{kji} + \alpha_{kj}(z_{kji} - \hat{M}_{kji}) & \text{if } i > 1 \\ z_{kji} & \text{if } i = 1 \end{cases} \\
\hat{M}_{kji+1} &= M_{kji} + V_{kji} \\
\hat{V}_{kji+1} &= V_{kji} \\
B_{kji+1} &\sim \text{Normal}(\hat{M}_{kji+1}, \lambda_{kj})
\end{aligned}$$

Figura 4-1: Representación jerárquica del modelo

en la condición k . Es importante notar que los parámetros e hiperparámetros relacionados con una variable de error, por ejemplo, las varianzas a nivel de condición y la incertidumbre interna, están definidas como sus inversos (denominados precisiones) y representados con la letra λ . Este ligero cambio de notación es necesario dado el programa implementado para realizar la estimación paramétrica, sin embargo, no tiene ningún efecto al inferir las variables de interés.

Capítulo 5

Resultados

5.1. Estimación del proceso generador

Como una medida general del desempeño de los participantes reportamos el error entre la media generadora (i.e. la posición “verdadera” de la nave) y las respuestas de los participantes. La Figura 5-1 muestra dichos valores agrupados para todos los participantes, en todos los ensayos, en cada una de las condiciones experimentales. En la parte inferior del gráfico se encuentran las frecuencias por condición representadas en histogramas, y el valor de su media y varianza. En todas las condiciones los valores más comunes se encuentran en cero o cerca de cero, lo cual implica que la mayoría de las veces no existía error, o era mínimo, entre las respuestas de los participantes y la media generadora. Dado que cada observación de la nave proviene de una distribución Normal, los valores más probables de aparición se encuentran cercanos a la media. Así, los participantes maximizaban su número de aciertos al elegir valores próximos a la media de la distribución, como se observa en la Figura 5-1. Sin embargo, es relevante notar que a medida que el nivel de ruido incrementa (siendo la condición uno la más ruidosa y la cuatro la menos ruidosa), también se incrementan los errores lejanos a cero, lo cual se aprecia en una varianza mayor para el conjunto de datos. Así mismo, puede apreciarse que los errores no se encuentran sesgados hacia valores positivos o negativos, lo cual implica que, en general, no existió una tendencia a subestimar o sobreestimar la posición. En la parte superior de la Figura 5-1 se encuentra una representación de la tarea y cómo se apreciaría el error de estimación de los participantes. La nave se encuentra centrada en el valor de la media y la franja

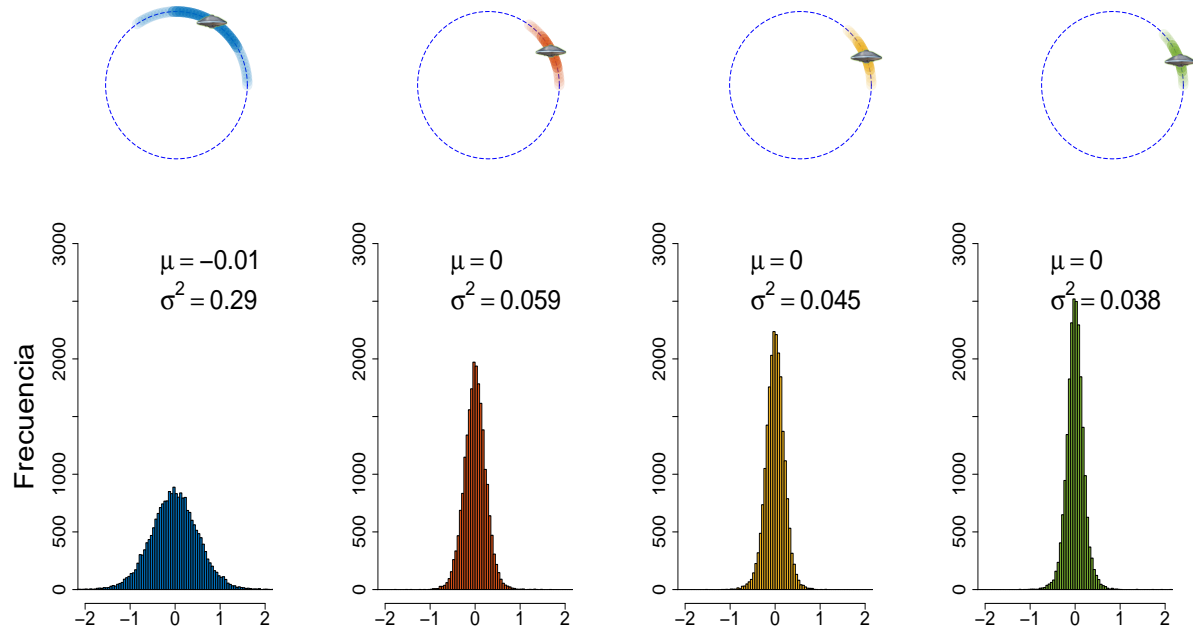


Figura 5-1: Errores agrupados. En la parte inferior se muestran histogramas de los errores entre la media generadora y las respuestas de los participantes para todos los ensayos de todos los participantes en cada condición (azul=condición 1, rojo= condición 2, amarillo=condición 3, verde=condición 4). En la parte superior se muestra cómo se apreciaría este error en la tarea.

de color representa el conjunto de errores agrupados. Con el fin de representar el desempeño individual de los participantes, en la figura 5-2 mostramos el error absoluto medio para cada uno de ellos en función de la razón $\frac{\sigma_x^2}{\sigma_z^2}$. Este valor nos permite conocer en promedio cuánto se alejaban de la posición verdadera de la nave por condición. Cada punto representa a un sujeto y aquellos de contorno oscuro representan las medias grupales. Es relevante destacar que, a diferencia de la Figura 5-1, los errores se graficaron sin importar si eran positivos o negativos, por este motivo las medias grupales, y las medias individuales, se alejan de 0. Se puede observar que, para la condición con la razón más pequeña, los errores absolutos son superiores al resto de las condiciones, lo cual es esperado dado que estaba perturbada por una mayor cantidad de ruido aleatorio. Para las demás condiciones, los errores absolutos son similares, lo cual implica que, aunque los participantes se encontraban bajo distintos niveles de ruido aleatorio, se desempeñaron de forma parecida.

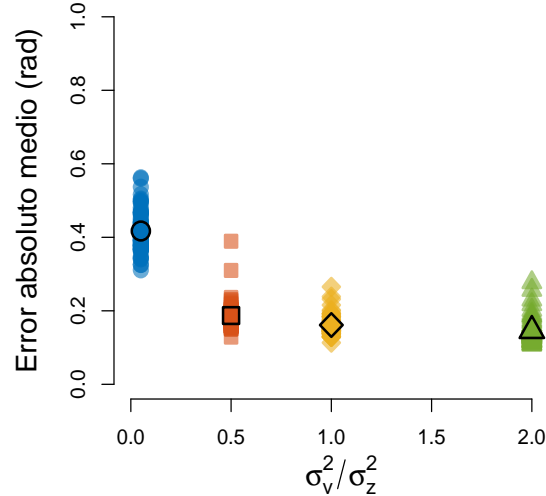


Figura 5-2: Error medio absoluto en función de $\frac{\sigma_v^2}{\sigma_z^2}$. Los puntos de contorno oscuro representan las medias grupales. Los puntos de contorno claro representan a cada participante.

5.2. Actualización ensayo a ensayo

En la Figura 5-3 mostramos el desempeño del participante 7 en la tarea. En general, el resto de los participantes mostraron un patrón similar. Para mayor claridad observamos únicamente 100 de los ensayos en cada condición (ensayo 100 al 200). Se puede apreciar que sus predicciones varían ensayo a ensayo siguiendo el valor de la media generadora (línea negra). En particular, cuando incrementa sus valores (i.e. la nave gira en dirección a las manecillas del reloj) o los disminuye (i.e. gira hacia el lado opuesto), el participante rastrea con precisión la dirección y la magnitud de los cambios. En la condición 1, se observa que en ocasiones las respuestas se alejan más de la media generadora en comparación con las demás condiciones, lo cual implica un mayor número de errores diferentes de cero como observamos para todos los participantes en las Figuras 5-1 y 5-2.

En resumen, las Figuras 5-1 y 5-2 nos muestran que los participantes predijeron adecuadamente la posición más probable de la nave en la mayor parte de los ensayos, y que no tenían una tendencia a subestimar o sobreestimar las posiciones. Así mismo, que la distancia que se alejaban de la media generadora fue mayor para la condición con más ruido, en comparación con las demás condiciones cuyos valores son similares. Finalmente, en la Figura 5-3, observamos

que las predicciones de los participantes varían ensayo a ensayo siguiendo los valores de la media generadora, análogo a un proceso de actualización vía ensayo y error.

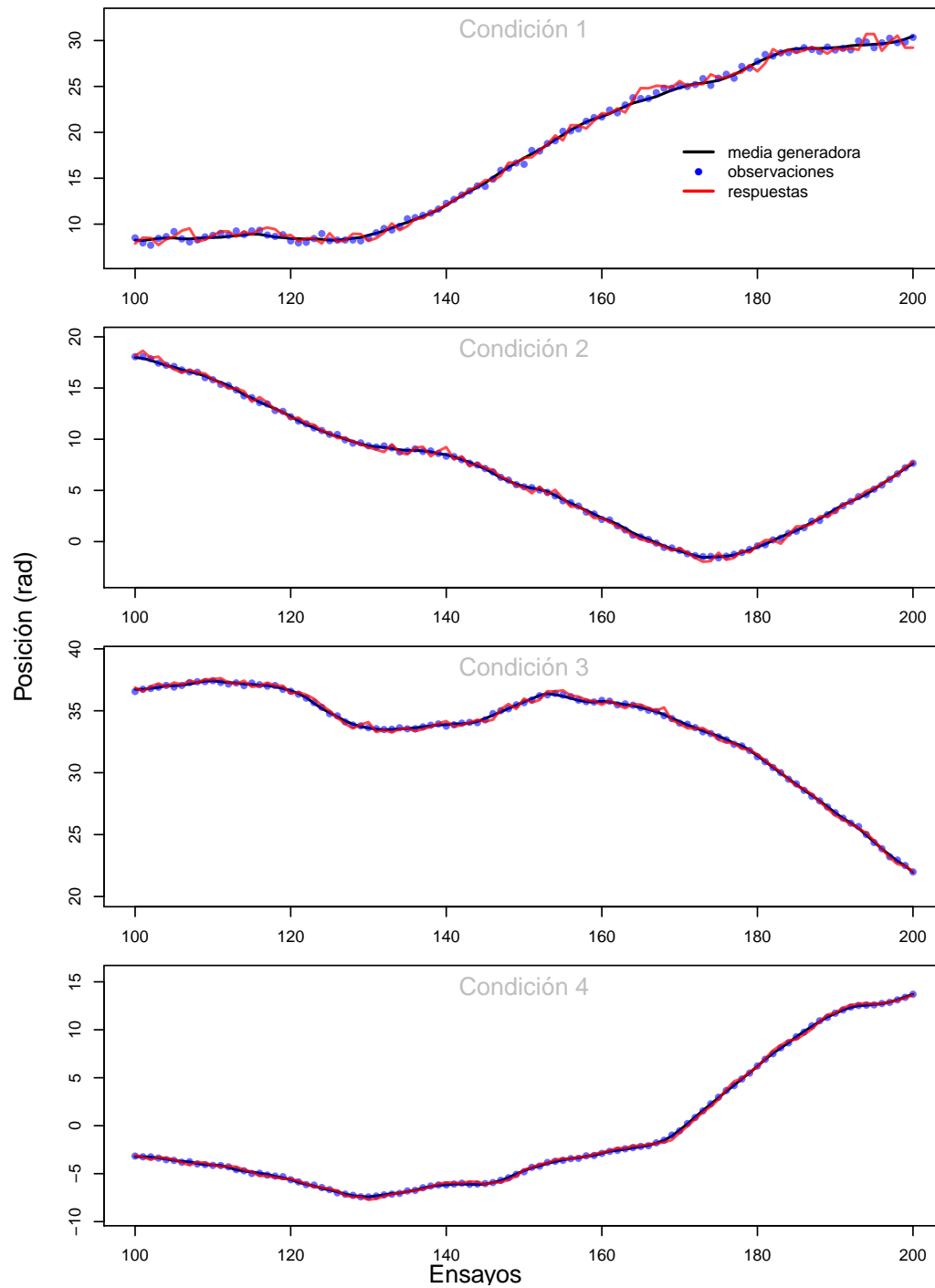


Figura 5-3: Observaciones y respuestas del participante 7. Se muestran 100 ensayos (100-200) en cada condición.

Capítulo 6

Estimación paramétrica

Una vez recabados los datos, el siguiente paso fue inferir los valores de los hiperparámetros de nuestro modelo representado en la Figura 4-1. Dado que conocemos B_{kji} , el vector de hiperparámetros \hat{H} , puede estimarse utilizando la regla de Bayes:

$$p(\hat{H} | B_{kji}) \propto p(B_{kji} | \hat{H})p(\hat{H}) \quad (6-1)$$

donde $p(\hat{H} | B_{kji})$ se denomina *distribución posterior* y representa las creencias sobre los valores de una vez observados los datos. $p(\hat{H})$ se conoce como *distribución prior*, y representa el conocimiento sobre los valores de \hat{H} antes de observar los datos. En la parte derecha de la Figura 4-1 se encuentran representadas las distribuciones prior para las medias por condición como Normales con media en 0 y precisión 0.001, lo cual indica que nuestro conocimiento sobre ellas antes de observar los datos es inespecífico. Para las precisiones por condición, utilizamos distribuciones prior Gamma, lo cual es una aproximación estándar en inferencia bayesiana (Lee y Wagenmakers, 2013, p. 56) dado su soporte para valores positivos. $(B_{kji} | \hat{H})$ se conoce como función de *verosimilitud* y permite estimar qué tan probable es que ocurran los datos considerando cada conjunto de valores de \hat{H} . Dado que asumimos que las respuestas de los participantes están distribuidas normalmente, la función de verosimilitud de nuestro modelo está definida por:

$$p(B_{kji} | \hat{H}) = \frac{1}{\sigma_{kj}\sqrt{2\pi}} e^{-\frac{(\hat{M}_{kji} - B_{kji})^2}{2\sigma_{kj}^2}} \quad (6-2)$$

Aproximamos las distribuciones posteriores de los hiperparámetros utilizando el software JAGS (Plummer, 2003), el cual implementa un método de muestreo conocido como Monte Carlo vía cadenas de Markov. Usamos tres cadenas independientes, cada una con un total de 12×10^4 muestreos, un período de burn-in (i.e. un conjunto de muestras que no son registradas con el fin de que el algoritmo pueda adaptarse) de 6×10^4 , y un thinning de 50 (i.e. se registraba uno de cada cincuenta valores de las cadenas). Se inspeccionó la convergencia de las cadenas mediante el estadístico \hat{R} (Gelman y Rubin, 1992), el cual compara la varianza entre cadenas con la varianza dentro de las cadenas, y valores cercanos a 1 indican convergencia. En general, se considera que valores de $\hat{R} > 1.1$ indican una convergencia dudosa. Ninguno de nuestros parámetros obtuvo un \hat{R} mayor a 1.01. En la Figura 6-1 presentamos una muestra con reemplazo de 150 valores

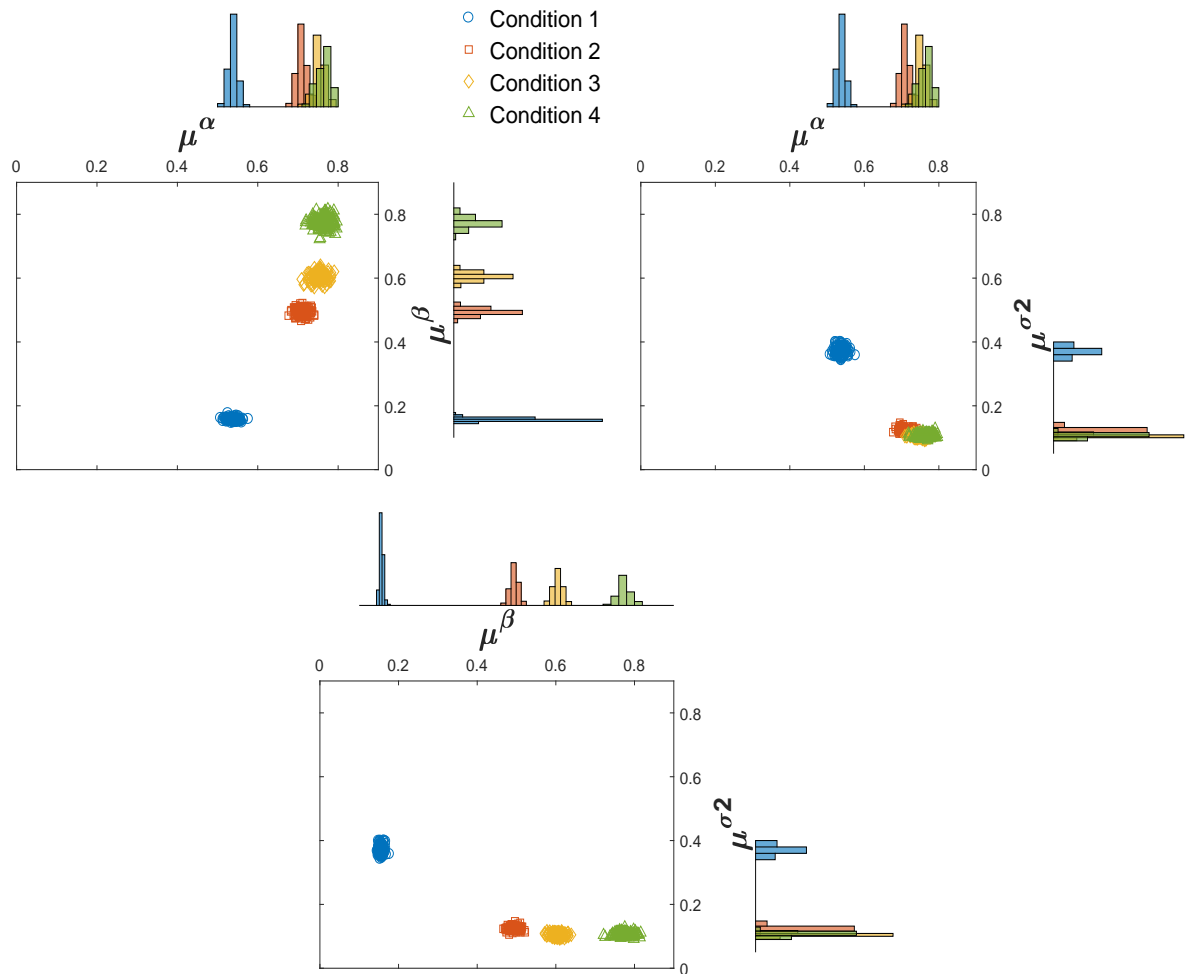


Figura 6-1: Distribuciones conjuntas posteriores para las medias por condición de las tasas de aprendizaje para posición (μ^α), velocidad (μ^β) e incertidumbre interna (μ^{σ^2}).

de las distribuciones posteriores conjuntas de las medias a nivel de condición. Este gráfico permite observar el comportamiento promedio de los parámetros de todos los participantes. En la parte superior izquierda, se observa μ^α y μ^β , correspondientes a las medias para las tasas de aprendizaje de la posición y velocidad, respectivamente, en cada condición. Se observa que conforme la condición se vuelve menos ruidosa, los participantes tienen a elevar ambas tasas de aprendizaje. Así mismo, notamos que los valores para μ^β parecen provenir de distribuciones claramente distintas, lo cual indica que los participantes respondían de manera diferencial al nivel de ruido en las observaciones para actualizar sus estimaciones sobre la velocidad de cambio del proceso generador. En el caso de μ^α , los valores para la condición 1 son los únicos claramente distinguibles de las demás condiciones. Sin embargo, se observa que para la condición 2, los valores se alejan ligeramente de la condición 3 y 4. Esto nos indica que al menos para la condición 1, los participantes implementaban valores distintos de α para actualizar sus estimados sobre la posición de la nave.

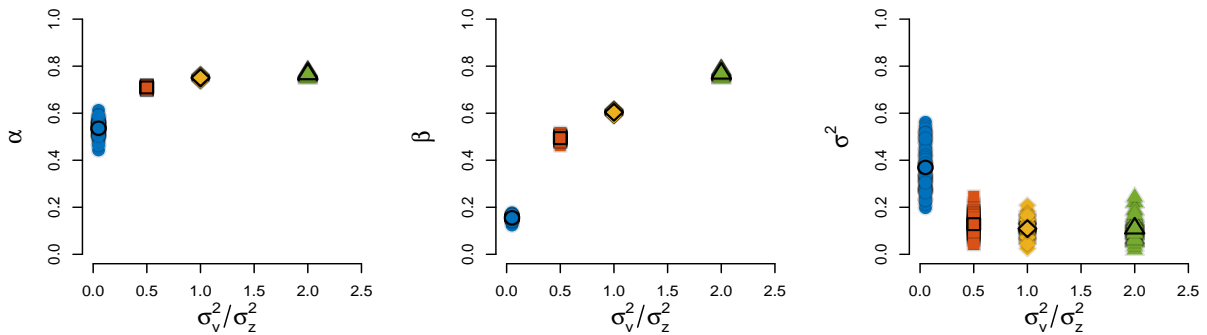


Figura 6-2: Máximos posteriores de los parámetros para cada sujeto. De izquierda a derecha se encuentran los valores para la tasa de aprendizaje para posición, velocidad e incertidumbre interna. Se observa que los tres parámetros presentan un comportamiento asintótico conforme la incertidumbre externa disminuye.

Los gráficos superior derecho e inferior, muestran el comportamiento de las tasas de aprendizaje y la incertidumbre interna (μ^{σ^2}). En general, observamos que, a medida que μ^{σ^2} disminuye, las tasas de aprendizaje tienen a elevarse tanto para el estimado de la posición como de la velocidad. En el caso de μ^{σ^2} , se observan valores similares entre las condiciones 2, 3 y 4, indicando que probablemente los valores de los parámetros de los participantes provienen de una misma

distribución, es decir, que para ese parámetro en tales condiciones, la manipulación experimental tuvo un efecto similar.

Aunque no se muestran, las distribuciones posteriores conjuntas de las varianzas permiten observar la variabilidad de las diferencias individuales. En su lugar, optamos por representar la inferencia sobre los parámetros individuales, lo cual nos permite apreciar de mejor manera dicha variabilidad. La Figura 6-2 muestra los valores de los máximos posteriores (i.e., aquellos más probables en las distribuciones posteriores de los parámetros) para cada participante en función de la razón $\frac{\sigma_v^2}{\sigma_z^2}$. Los puntos de contorno oscuro representan las medias grupales y los de contorno claro a cada sujeto. El panel izquierdo y central muestran que las tasas de aprendizaje tanto para posición (α) como velocidad (β) tienden a incrementarse de manera asintótica conforme disminuye la incertidumbre externa (i.e. la razón $\frac{\sigma_v^2}{\sigma_z^2}$ incrementa). En el panel derecho, observamos que en el caso de la incertidumbre interna, sus valores disminuyen asintóticamente conforme disminuye la incertidumbre externa.

Capítulo 7

Precisión del modelo

Una vez realizada la inferencia sobre los parámetros e hiperparámetros del modelo, la siguiente pregunta es ¿qué tan bien describe el comportamiento de los participantes? Es importante recordar que en cada ensayo el modelo calcula una distribución de la posición de la nave. En particular, una distribución Normal con media \hat{M}_{kji} y varianza σ_{kj}^2 . Para estimar su precisión, computamos la diferencia entre los valores de \hat{M}_{kji} y las respuestas de los participantes. Valores cercanos a 0, indican que la respuesta está cerca de la posición más probable predicha por el modelo, es decir, \hat{M}_{kji} . Así mismo, a medida que los valores se alejan de cero, la probabilidad de que la predicción corresponda con la respuesta va disminuyendo. Dado que en cada condición k , para el participante j existe un valor distinto del parámetro σ^2 (incertidumbre interna), una forma de representar las discrepancias entre las predicciones del modelo y las respuestas para todos los participantes, es estandarizarlas por σ_{kj} . Para ello, seguimos la siguiente ecuación:

$$error_{kji} = \frac{\hat{M}_{kji} - B_{kji}}{\sigma_{kj}} \quad (7-1)$$

De esta manera, podemos representar los errores del modelo sobre una distribución Normal estándar. La Figura 7-1 muestra dichos valores para todos los sujetos en las cuatro condiciones. La franja en color gris muestra el rango de predicción del modelo $\pm\sigma$ y $\pm 2\sigma$, y la distribución a un costado izquierdo es la predicha por el modelo en cada ensayo, cuya media en 0 representa el valor de \hat{M}_{kji} . En general, mientras los valores sean más próximos a 0, mejor es la descripción del modelo sobre las respuestas de ese participante. Como podemos observar, en todas las

condiciones el desempeño de los participantes es capturado por el modelo, es decir que en cada ensayo, las respuestas se encontraban en el lugar de mayor densidad de la distribución predicha.

Con el fin de mostrar el desempeño del modelo ensayo a ensayo, en la Figura 7-2 presentamos la secuencia de respuestas del participante 7 junto con las predicciones del modelo. Se realizaron simulaciones con una muestra de 150 valores de la distribución posterior conjunta de los parámetros del participante. La línea negra indica las respuestas y la franja gris el conjunto de predicciones realizadas con la muestra de valores de los parámetros. Como se puede observar, la secuencia de respuestas del participante en cada condición es capturada por el modelo.

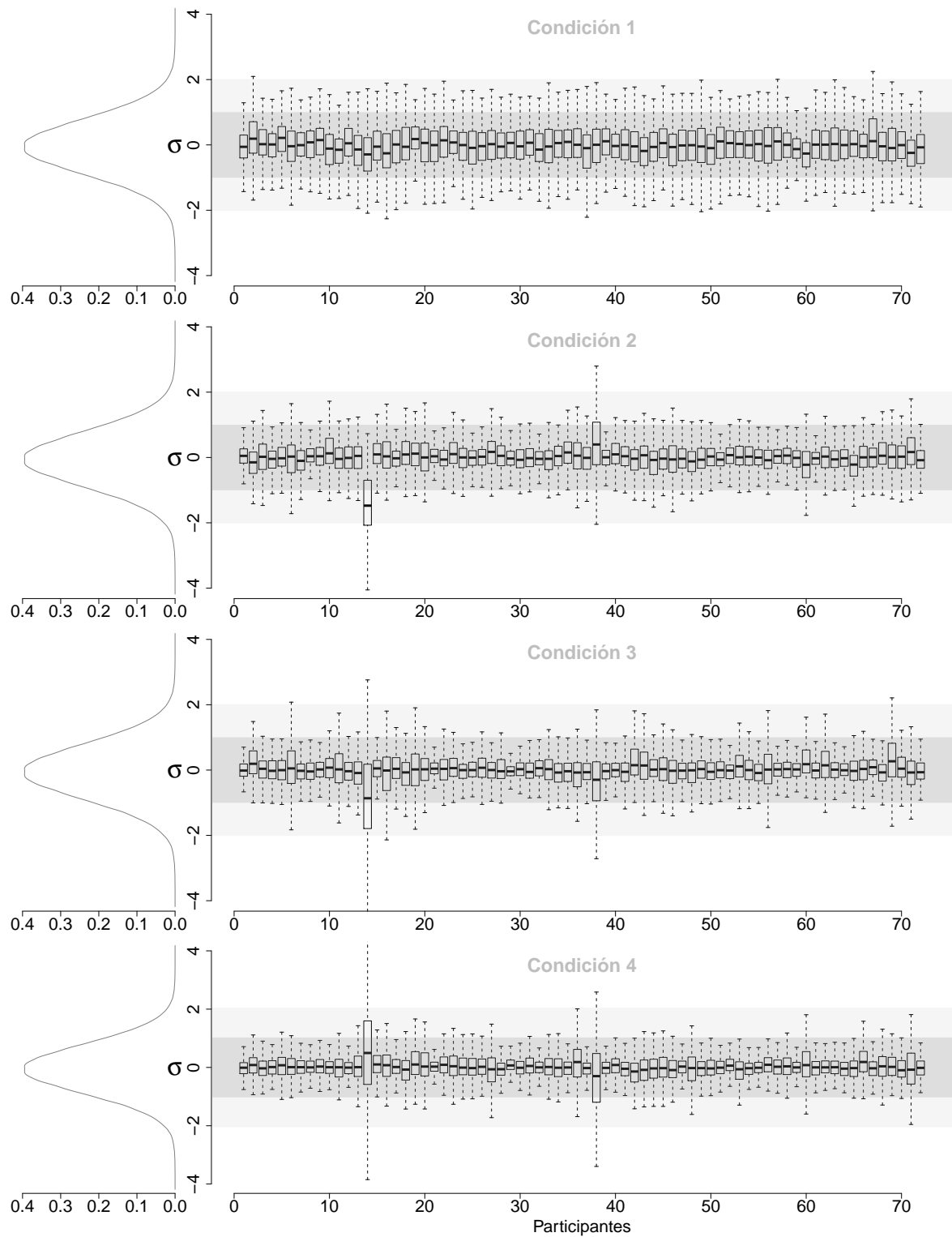


Figura 7-1: Errores estandarizados entre las predicciones del modelo y las respuestas de los sujetos

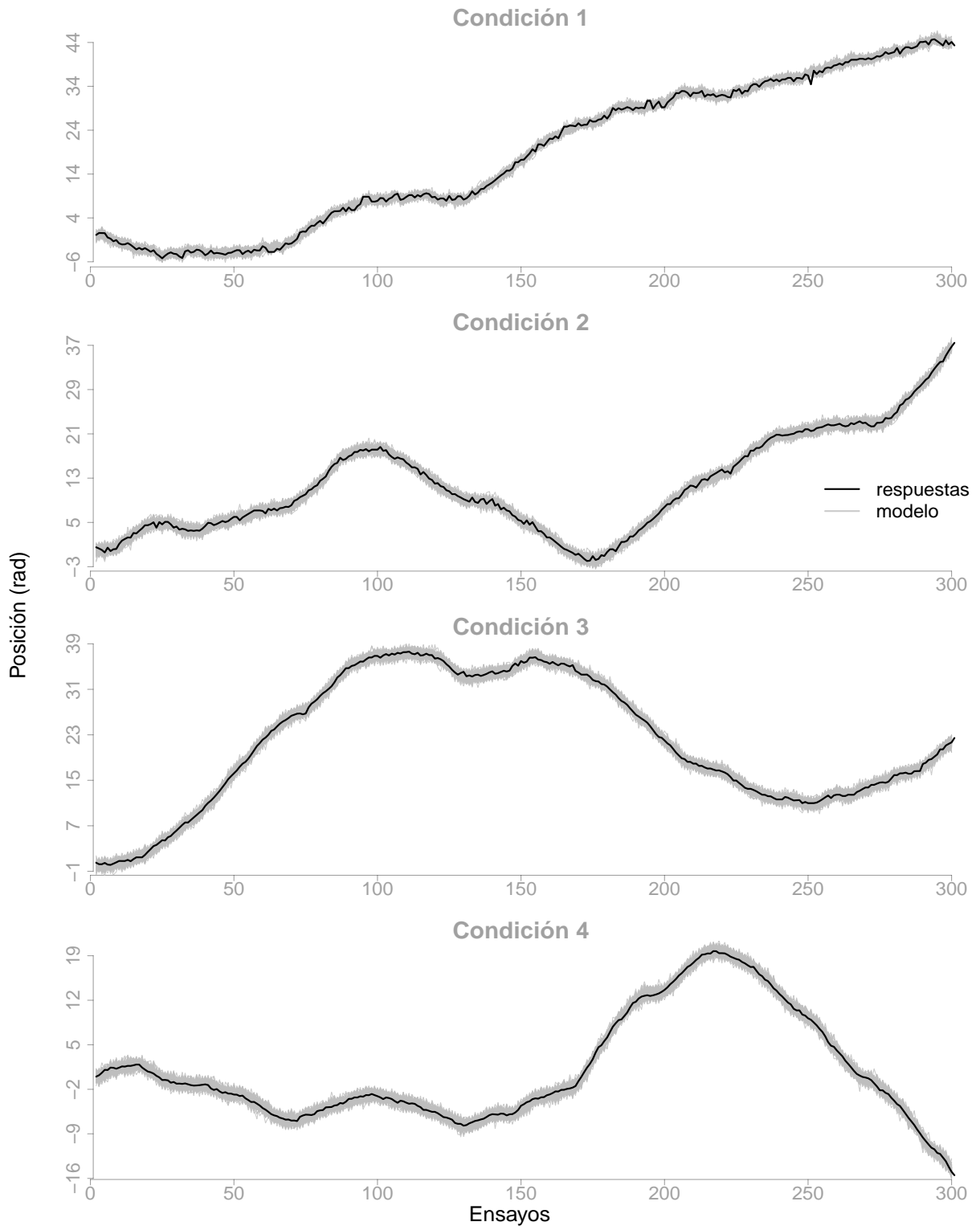


Figura 7-2: Respuestas y predicciones del modelo para el participante 7.

Capítulo 8

Discusión

En el presente trabajo estudiamos cómo las personas hacen predicciones en un entorno que se modifica momento a momento y del cual solo tienen observaciones ruidosas. En particular, nuestros participantes estimaron la posición de una nave espacial que variaba en función de un término de velocidad y ruido aleatorio. Los resultados muestran que la discrepancia entre las predicciones de los participantes y la media generadora de la nave es cero o cercana a cero en la mayor parte de los ensayos, lo cual indica que son capaces de estimar la posición más probable de la nave en el ensayo siguiente. Un modelo jerárquico que implementa una regla de aprendizaje derivada del Filtro de Kalman es capaz de describir su comportamiento de manera adecuada como se muestra en las Figuras 7-1 y 7-2. La estimación paramétrica sugiere que conforme disminuye la incertidumbre externa los participantes incrementan los valores de sus tasas de aprendizaje y disminuyen su incertidumbre interna, lo cual indica que confían cada vez más en las observaciones para actualizar sus estimaciones. En particular, los valores de las tasas de aprendizaje para velocidad parecen provenir de distribuciones distintas en cada condición, lo cual se observa en la Figura 6-1 para el hiperparámetro μ^β . Esto indica que la actualización de la velocidad es sensible a las variaciones de incertidumbre externa entre condiciones. Los valores para las tasas de aprendizaje de posición y la incertidumbre interna parecen provenir de una misma distribución en la condición 3 y 4, lo cual refleja que ambas alcanzan un comportamiento asintótico antes que las tasas de aprendizaje para velocidad como se observa en la Figura 6-2.

8.1. Tasas de aprendizaje e incertidumbre interna asintóticas

De forma interesante, observamos que ambas tasas de aprendizaje y la incertidumbre interna de los participantes muestran un comportamiento asintótico conforme disminuye la incertidumbre externa (Ver Figura 6-2). Teóricamente, es un patrón importante. En el caso de las tasas de aprendizaje, su relación con la incertidumbre externa de acuerdo al Filtro de Kalman está dada por:

$$K = \frac{\sigma_{INT}^2}{\sigma_{INT}^2 + \sigma_{EXT}^2} \quad (8-1)$$

donde K es una matriz denominada Ganancia de Kalman, la cual contiene los valores de las tasas de aprendizaje para posición y velocidad; σ_{INT}^2 representa la incertidumbre interna y σ_{EXT}^2 la incertidumbre externa. Es decir, las tasas de aprendizaje son iguales a la incertidumbre relativa (Daw, 2013) (i.e., la incertidumbre interna entre la incertidumbre total). Si graficamos la parte derecha de la ecuación 8-1 en función de una razón señal-ruido, tal como lo hicimos con los participantes, obtenemos asíntotas similares a las observadas en la Figura 6-2. Es decir, una posibilidad es que los participantes modulen el impacto de las nuevas observaciones computando la incertidumbre relativa, tal como se especifica en el Filtro de Kalman, y que nosotros hemos reportado simplemente como tasas de aprendizaje.

De forma análoga, la incertidumbre teóricamente interna se actualiza mediante:

$$\sigma_{INT_t}^2 = (1 - K)\sigma_{INT_{t-1}}^2 \quad (8-2)$$

Si graficamos $(1 - K)\sigma_{INT_{t-1}}^2$ en función de una razón señal-ruido, obtenemos una asíntota similar a la observada en el panel derecho de la Figura 6-2. En otras palabras, nuestro modelo estima los valores de $(1 - K)\sigma_{INT_{t-1}}^2$, los cuales implican un cálculo de acuerdo al algoritmo de donde provienen. Si asumimos que los supuestos del modelo son verdaderos, los participantes estarían computando las ecuaciones 8-1 y 8-2.

8.2. Valores óptimos de los parámetros

Aunque es probable que los participantes implementen un algoritmo de predicción similar al propuesto en este trabajo, una pregunta relevante es si lo hacen de manera óptima. Con el fin de

indagar en esta cuestión, dejamos que el modelo hiciera sus propias predicciones sobre el proceso generador. Es decir, lo pusimos a jugar la misma tarea que a los sujetos. Su comportamiento puede considerarse como óptimo en la medida en que su tarea consiste en inferir los parámetros verdaderos que subyacen al experimento. Si estos valores corresponden con los inferidos para los sujetos, podemos decir que estos últimos realizan predicciones óptimas. En la Figura 8-1 observamos los valores de los máximos posteriores para los parámetros de cada sujeto tal y como se mostró en la Figura 6-1.

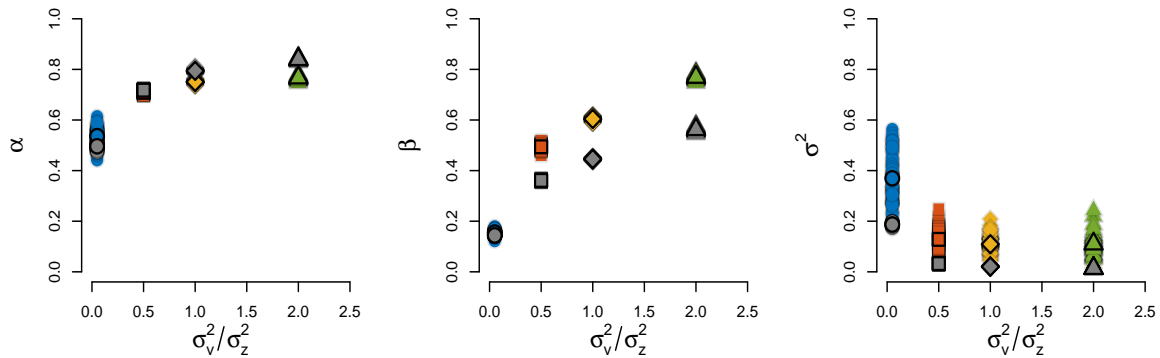


Figura 8-1: Máximos posteriores de los parámetros inferidos para los participantes y para el modelo óptimo.

En gris se encuentran los máximos posteriores para cada “sujeto óptimo”. Los puntos de contorno oscuro representan las medias grupales. Podemos notar que las tasas de aprendizaje para posición presentan valores cercanos a aquellos predichos por el modelo óptimo en las cuatro condiciones. Sin embargo, para las tasas de aprendizaje de velocidad, esto ocurre solamente en la condición con mayor incertidumbre externa (azul). En las demás, los valores para los sujetos son superiores a los del modelo óptimo, lo cual indica que sus estimaciones de la velocidad cambiaban con más frecuencia conforme se obtienen más observaciones. Por otro lado, los valores para la incertidumbre interna de los sujetos son mayores en todas las condiciones que los del modelo óptimo. Esto indica que las estimaciones de los participantes en general presentaban una mayor variabilidad. Los máximos posteriores para cada uno de los tres parámetros del modelo óptimo son también aquellos que minimizan el error de predicción con las observaciones. Por ello, cualquier desviación implica a su vez un mayor grado de error por parte de los participantes al inferir la posición verdadera de la nave con respecto del desempeño óptimo.

8.3. Actualización ensayo a ensayo

Actualmente existe un continuo debate sobre las reglas de aprendizaje que implementan las personas y otros animales para actualizar sus creencias en entornos cambiantes. Los modelos de reforzamiento, como el propuesto en este trabajo, sugieren que este proceso se lleva a cabo ensayo a ensayo vía reducción del error. Existe evidencia de que el proceso de actualización podría ocurrir de forma distinta. Gallistel et al. (2014) y Ricci y Gallistel (2017) sugieren que al estimar probabilidades las personas actualizan sus creencias solo cuando existe suficiente evidencia para hacerlo. Un proceso de esta naturaleza implica que los participantes podrían mantener una misma estimación por varios ensayos y cambiarla abruptamente a un nuevo valor, produciendo un patrón escalonado como el reportado en sus estudios. Nuestros resultados muestran que, al menos al estimar posiciones, los participantes parecen actualizar sus estimaciones ensayo a ensayo y no de manera abrupta (ver Figura 5-3). Esta diferencia puede deberse al grado en que los cambios de la variable de interés pueden discriminarse. Cuando las variaciones son graduales, en el caso de probabilidad podría tomar una mayor cantidad de ensayos hasta que el cambio es finalmente detectado, y entonces ocurre un ajuste. Es decir, a pesar de que los valores se modifican continuamente, es difícil para las personas detectar estos cambios ensayo a ensayo. Esta tarea es más sencilla cuando se estiman posiciones, dado que es una variable que los participantes pueden observar de manera directa, y de esta manera pueden hacer ajustes a sus estimaciones en cada ensayo.

8.4. Implicaciones para los modelos de ensayo y error

El modelo tradicional de reforzamiento (Rescorla y Wagner, 1972), así como sus versiones generalizadas (Sutton y Barto, 1981) han permitido realizar una gran variedad de predicciones de fenómenos conductuales (Daw, 2013; Daw y Tobler, 2013). Así mismo, evidencia en Neurociencia, indica que es un modelo biológicamente plausible, sugiriendo que el cómputo de los errores de predicción (parte central de este modelo) podría llevarse a cabo en neuronas dopaminérgicas del mesencéfalo (Schultz, 1997). Una parte importante de la evidencia neural de este modelo es que el error de predicción es una señal global que se dirige a múltiples regiones del cerebro, pudiendo tener funciones distintas en cada una (Daw, 2013). En este trabajo, nuestro modelo

sugiere que el error de predicción podría implementarse para actualizar variables que controlan la dinámica del entorno. En el caso más sencillo, la velocidad a la que se mueve, es decir una primera tasa de cambio. Sin embargo, sería importante estudiar si las personas y otros animales pueden hacer estimaciones adecuadas cuando el entorno se modifica por una tasa de la tasa. En el caso en que la variable es posición, nos referimos a la aceleración. Una versión generalizada del Filtro de Kalman permite actualizar las n posibles tasas que modifican una variable vía el error de predicción. Estudios futuros podrían indagar si es un modelo apropiado de aprendizaje cuando las personas se encuentran en entornos de esta complejidad. Así mismo, en este trabajo los participantes realizaron predicciones en una sola dimensión, la cual se representó en radianes. En entornos naturales deben hacerlo en un espacio tridimensional. Para ello, la posición, velocidad y aceleración deben estimarse para el eje de las X, Y y Z o sus análogos en coordenadas esféricas. Experimentos futuros pueden explorar la tarea de predicción de posiciones en más de una dimensión. El Filtro de Kalman es un modelo que permite extender su gama de predicciones a este tipo de entornos.

Capítulo 9

Conclusiones

En este trabajo mostramos que nuestros participantes pueden inferir de manera correcta la posición más probable de una nave espacial cuando esta varía en función de un término de velocidad y de ruido aleatorio o incertidumbre externa. Una versión generalizada del modelo clásico de reforzamiento que estima la velocidad de cambio del entorno es capaz de describir los datos. La inferencia paramétrica indica que a medida que la incertidumbre externa disminuye, hay un incremento de las tasas de aprendizaje y una disminución de la incertidumbre interna, indicando que los participantes asignan una mayor importancia a sus observaciones para actualizar sus estimaciones. En ambos casos, este patrón ocurre de manera asintótica, sugiriendo un posible cómputo de la incertidumbre relativa del entorno, como se especifica teóricamente en el Filtro de Kalman. Por otro lado, mostramos que los participantes presentan estimaciones sobre la velocidad más inestables y una mayor incertidumbre interna que un modelo óptimo, lo cual indica que a pesar de rastrear la dinámica de la tarea, presentan una mayor cantidad de errores que si conocieran de manera perfecta la estructura del experimento. En general, este trabajo provee evidencia de que los algoritmos de reforzamiento vía ensayo y error son mecanismos plausibles de aprendizaje en entornos cambiantes. En particular, el modelo implementado sugiere que cuando los cambios ocurren momento a momento, las personas podrían estimar variables de segundo orden como la velocidad de cambio e integrarlas a sus predicciones.

Capítulo 10

Referencias

- Baum, W. M. (2010). Dynamics of choice: A tutorial. *Journal of the Experimental Analysis of Behavior*, *94*(2), 161-174.
- Bouzas (2017). Adaptación a Propiedades Estadísticas de Sucesos Biológicamente Importantes. En J.A. Camacho, F. Cabrera, O. Zamora, F.H. Martínez y J.J. Irigoyen (Coord) *Aproximaciones al Estudio del Comportamiento y sus Aplicaciones, Vol I* (pp. 13-35) Universidad Autónoma de Tlaxcala.
- Brainard, D. H. (1997) The Psychophysics Toolbox. *Spatial Vision* *10*:433-436.
- Bush, R. R., y Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*(5), 313-323.
- Daw, N. D. (2013). Advanced reinforcement learning. *Neuroeconomics: Decision Making and the Brain*, 299-320.
- Daw, N. D., y Tobler, P. N. (2013). Value learning through reinforcement: the basics of dopamine and reinforcement learning. *Neuroeconomics*, 283-298.
- Gallistel, C. R., Krishan, M., Liu, Y., Miller, R., y Latham, P. E. (2014). The perception of probability. *Psychological Review*, *121*(1), 96.
- Gallistel, C. R., Mark, T. A., King, A. P., y Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*(4), 354.
- Geist, M., Pietquin, O., y Fricout, G. (2008). Bayesian reward filtering. In *European Workshop on Reinforcement Learning* (pp. 96-109). Springer, Berlin, Heidelberg.

- Gelman, A., y Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 457-472.
- Gershman, S. J. (2015). A unifying probabilistic view of associative learning. *PLoS Computational Biology*, 11(11), e1004567.
- Gershman, S. J. (2017). Dopamine, Inference, and Uncertainty. *bioRxiv*, 149849.
- Grewal, Mohinder S., y Angus P. Andrews (2014). Kalman filtering: Theory and Practice with MATLAB. *John Wiley & Sons*, 1:25
- Innis, N. K., y Staddon, J. E. R. (1971). Temporal tracking on cyclic-interval reinforcement schedules. *Journal of the Experimental Analysis of Behavior*, 16(3), 411-423.
- Kakade, S., y Dayan, P. (2000). Acquisition in autoshaping. In S. A. Solla, T. K. Leen, & K.-R. Müller (Eds.), *Advances in neural information processing*, 12 (pp. 24–30). Cambridge, MA: MIT Press.
- Kakade, S., y Dayan, P. (2002). Acquisition and extinction in autoshaping. *Psychological Review*, 109(3), 533.
- Kalman, R. E.(1960) . A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82.1: 35-45.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., y Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1.
- Lee, M. D., y Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.
- Lee, M.D. (2016). Bayesian methods in cognitive modeling. Submitted for *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*. Fourth Edition.
- Miller, R. R., Barnet, R. C., y Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin*, 117(3), 363.
- Nassar, M. R., Wilson, R. C., Heasly, B., y Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, 30(37), 12366-12378.
- Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision* 10:437-442.

- Plummer, M. (2003, March). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *In Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, p. 125).
- Rao, R., y Ballard, D. (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Computation*, 9, 721–763.
- Rao, R., y Ballard, D. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2, 79–87.
- Rescorla R.A. y Wagner A.R. (1972) A Theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Black AH, Prokasy WF, editors, Classical conditioning II: current research and theory*. New York: Appleton Century Crofts. chapter 3. pp. 64–99.
- Ricci, M., y Gallistel, R. (2017). Accurate step-hold tracking of smoothly varying periodic and aperiodic probability. *Attention, Perception, & Psychophysics*, 1-15.
- Schultz, W., Dayan, P., y Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593-1599.
- Shiffrin, R. M., Lee, M. D., Kim, W., y Wagenmakers, E. J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science*, 32(8), 1248-1284.
- Speekenbrink, M., y Shanks, D. R. (2010). Learning in a changing environment. *Journal of Experimental Psychology: General*, 139(2), 266.
- Sutton, R. S. (1992, May). Gain adaptation beats least squares. *In Proceedings of the 7th Yale workshop on adaptive and learning systems* (Vol. 161168).
- Sutton, R. S., y Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review*, 88(2), 135.
- Sutton, R. S., y Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1). Cambridge: MIT press.
- Wilson, R. C., Nassar, M. R., y Gold, J. I. (2013). A mixture of delta-rules approximation to Bayesian inference in change-point problems. *PLoS Computational Biology*, 9(7), e1003150.