

# Trabajo Estadística Bayesiana

Eugenio Guzmán\*    Alejandra Molina†    Jaquelin Morillo‡    Diego Ramirez§  
Francisco Villarroel¶

2022-08-26

## Introducción

La reciente literatura neurocientífica ha demostrado que los procesos mentales involucrados en la toma de decisiones pueden ser modelados matemáticamente. En particular la distribución de los tiempos de decisión es concordante con una distribución de un Wiener Drift Diffusion Model. En este trabajo se proponen 5 posibles modelos de toma de decisión basados en el DDM para intentar explicar los resultados de un experimento de laboratorio donde los sujetos tuvieron que escoger entre dos alternativas. Los parámetros de los modelos se han calculado utilizando una metodología Bayesiana y se han evaluado los desempeños de esos modelos con los criterios de información de devianza (DIC) y con la validación “Leave-one-out” (LOO).

## Drift Diffusion model

El *Drift Diffusion Model* (DDM) es un modelo teórico de tiempos de respuesta para la toma de decisiones entre dos alternativas. El modelo consiste en la acumulación secuencial de evidencia a favor de una u otra alternativa hasta que se cruza un umbral hacia una de las opciones, momento en que la decisión es tomada. Entre el momento que se expone el sujeto a las alterantivas y el momento en que toma la decisión, la deriva mental entre una opción y otra es el “drift” o tasa de convergencia hacia la alternativa que es percibida como correcta.

El modelo entrega como resultado un valor neto que corresponde al tiempo total en la toma de la decisión y un signo (positivo o negativo) que indica la alternativa elegida (en el caso de este trabajo, si se eligió la opción de la izquierda o de la derecha).

En este trabajo se ha utilizado el modelo DDM Wiener que tiene como parámetros: el tiempo de no-decisión ( $\tau$ ), que corresponde al tiempo de procesamiento mental de la tarea en actividades no relacionadas a la decisión en sí misma; la separación entre las alternativas ( $\alpha$ ), que es la diferencia percibida entre las alternativas y da cuenta de la rapidez versus precisión en la toma de la decisión; el sesgo ( $\beta$ ): que corresponde a una tendencia no explicada por los parámetros del experimento que lleva a la persona a decidir más por una opción que por la otra; el “drift” ( $\delta$ ), que corresponde a la tasa a la que se converge hacia una de las alternativas, este parámetro da cuenta de la ambigüedad en el experimento, mientras mayor es el drift mayor es la ambigüedad en la toma de decisión y por lo tanto mayor el tiempo que se destina a elegir una de las alternativas.

---

\*ejguzmanl@udd.cl

†alejandramm@gmail.com

‡jaquelin.morillo@gmail.com

§diegoramirez.al34@gmail.com

¶fvillarroelr@udd.cl

## Explicación del experimento

Veinte personas participaron en un juego de ruleta, donde debían apostar a uno de los dos colores: naranja o celeste. Cada color está asociado a una probabilidad de dar un premio, el que está indicado por el área de ese color en la barra de una ruleta, y al monto de dicho premio potencial, indicado por el número sobre la barra. Además, algunas ruletas (botones de la barra) tenían áreas grises que representan una máscara que oculta el verdadero color de esas áreas. Esta condición se conoce como ambigüedad, y tuvo un rango de 0 (sin ambigüedad); 0.3; 0.4; 0.5; 0.6. Cada persona realizó el experimento en varias oportunidades obteniéndose un total de 3000 muestras.

Las variables presentes en la base de datos y que se utilizaron para los cálculos son:

- **P**: Probabilidad del color naranja. La probabilidad del color celeste es el complemento (1-p).
- **O**: Premio normalizado del color naranja. El premio del color celeste es el complemento (1-o)
- **a**: Rango de la máscara o grado de ambigüedad  $a=[0 .3 .4 .5 .6 ]$
- **bias** : Lugar donde estaba la verdadera división de los colores dentro de la máscara. 0: margen izquierdo, 0.5: mitad, 1: margen derecho.
- **pv**: Probabilidad visible izquierdo, ignorando el área de la máscara. probabilidad visible derecha 1-pv.
- **ttp**: Tiempo de aparición de las opciones, en sub milisegundos. 10000 sub milisegundos = 1 segundo.
- **ttr**: Tiempo de la respuesta.
- **F**: Resultados (1: izquierda, 2 Derecha)
- **r**: Decisión (1: izquierda, 2 Derecha)
- **B**: Bloque o repetición del experimento en el mismo sujeto
- **sub**: Identificador de cada participante

## Metodología

El *Drift Diffusion Model* (DDM) es un modelo teórico de tiempos de respuesta para la toma de decisiones entre dos alternativas. El modelo consiste en la acumulación secuencial de evidencia a favor de una u otra alternativa hasta que se cruza un umbral hacia una de las opciones, momento en que la decisión es tomada. Entre el momento que se expone el sujeto a las alternativas y el momento en que toma la decisión, la deriva mental entre una opción y otra es el “drift” o tasa de convergencia hacia la alternativa que es percibida como correcta.

El modelo entrega como resultado un valor neto que corresponde al tiempo total en la toma de la decisión y un signo (positivo o negativo) que indica la alternativa elegida (en el caso de este trabajo, si se eligió la opción de la izquierda o de la derecha).

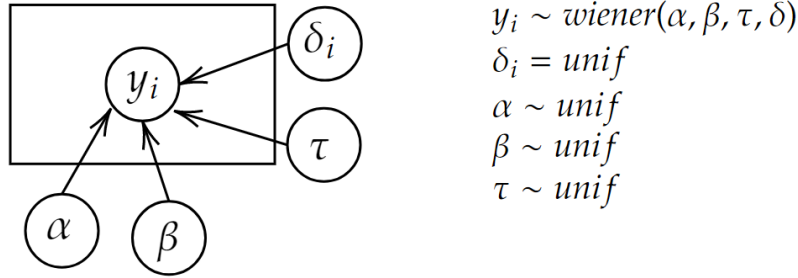
En este trabajo se ha utilizado el modelo DDM Wiener que tiene como parámetros: el tiempo de no-decisión ( $\tau$ ), que corresponde al tiempo de procesamiento mental de la tarea en actividades no relacionadas a la decisión en sí misma; la separación entre las alternativas ( $\alpha$ ), que es la diferencia percibida entre las alternativas y da cuenta de la rapidez versus precisión en la toma de la decisión; el sesgo ( $\beta$ ): que corresponde a una tendencia no explicada por los parámetros del experimento que lleva a la persona a decidir más por una opción que por la otra; el “drift” ( $\delta$ ), que corresponde a la tasa a la que se converge hacia una de las alternativas, este parámetro da cuenta de la ambigüedad en el experimento, mientras mayor es el drift mayor es la ambigüedad en la toma de decisión y por lo tanto mayor el tiempo que se destina a elegir una de las alternativas.

| Symbol   | Parameter           | Interpretation   |
|----------|---------------------|--|
| $\alpha$ | Boundary separation | Speed-accuracy trade-off<br>(high $\alpha$ means high accuracy)              |
| $\beta$  | Initial bias        | Bias for either response<br>( $\beta > 0.5$ means bias towards response 'A') |
| $\delta$ | Drift rate          | Quality of the stimulus<br>(close to 0 means ambiguous stimulus)             |
| $\tau$   | Nondecision time    | Motor response time, encoding time<br>(high means slow encoding, execution)  |

Para modelar la toma de decisión en este experimento se proponen 5 modelos basados en el DDM explicado anteriormente, en donde se propone que el parámetro delta (la tasa de convergencia o “drift”) puede estar relacionado con la información que tiene el sujeto sobre las alternativas que se le dan.

### Modelo 1:

Delta es un parámetro aleatorio que proviene a priori de una distribución uniforme, al igual que los otros parámetros del modelo.



### Modelos 2 al 5

Delta es una función lineal de los tamaños de las barras (p), que equivale a la probabilidad de que esa barra gane en la ruleta y de los montos de los premios (o) y los demás parámetros se consideraron a priori provenientes de una distribución uniforme.

#### Modelo 2

$$\delta = b_0 + b_1 p + b_2 o$$

#### Modelo 3

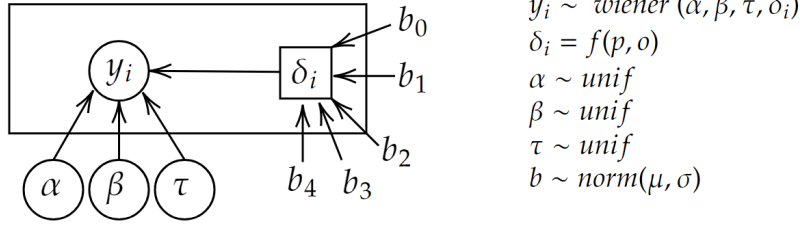
$$\delta = b_0 + b_1 p + b_2 o + b_3 p^2 + b_4 o^2$$

#### Modelo 4

$$\delta = b_0 + b_1 o + b_2 p^2$$

#### Modelo 5

$$\delta = b_0 + b_1 o + b_2 o^2$$



Para calcular los parámetros de los distintos modelos se utilizó la metodología Bayesiana con la función de enlace Wiener, a través del módulo JAGS Wiener (Wabersich & Vandekerckhove, 2014) en JAGS y el paquete rjags en R-statistics. Los resultados de las distribuciones de probabilidad de los parámetros de todos los modelos, el análisis de estos parámetros y los análisis de los desempeños de los modelos se detallan en las siguientes secciones del informe.

## Resultados - Modelos y tablas sumarias

### Modelo 1

$$Y_{(ij)} \sim \text{Wiener}(\alpha_{(ij)}, \beta_{(ij)}, \tau_{(ij)}, \delta_{(ij)})$$

Este modelo supone una distribución Wiener dependiente de los parámetros alfa, beta, tau y delta en función de los sujetos estudiados y los *trials*.

Table 1: Media de los parámetros modelados

| Modelo 1  |           |           |          |          |          |
|-----------|-----------|-----------|----------|----------|----------|
| Parámetro | A         | B         | C        | D        | E        |
| $\alpha$  | 0.76461   | 0.9084    | 0.92244  | 0.93057  | 0.95987  |
| $\beta$   | 0.46303   | 0.51041   | 0.49342  | 0.5304   | 0.54067  |
| $\tau$    | 0.0045092 | 0.0081149 | 0.001054 | 0.017666 | 0.016152 |
| $\delta$  | 0.66599   | 0.085327  | 0.36738  | -0.13167 | 0.26716  |

Figure 1: Tabla de parámetros del modelo

### Modelo 2

$$Y_{(ij)} \sim \text{Wiener}(\alpha_{(ij)}, \beta_{(ij)}, \tau_{(ij)}, \delta_{(p_{ij}, o_{ij})})$$

Al igual que el modelo 1 se contemplan los mismos parámetros, con la excepción que el parámetro delta depende linealmente de la probabilidad y del premio normalizado

Table 2: Media de los parámetros modelados

| Modelo 2  |           |           |            |          |          |
|-----------|-----------|-----------|------------|----------|----------|
| Parámetro | A         | B         | C          | D        | E        |
| $\alpha$  | 0.88527   | 0.9671    | 0.96805    | 0.95355  | 0.9938   |
| $\beta$   | 0.45883   | 0.502     | 0.49792    | 0.52661  | 0.53689  |
| $\tau$    | 0.0021536 | 0.0061823 | 0.00085613 | 0.016333 | 0.015132 |
| $b_0$     | 11.642    | 6.8589    | 5.3276     | 3.0593   | 2.7807   |
| $b_1$     | -13.817   | -7.7817   | -5.7075    | -2.9145  | -1.3458  |
| $b_2$     | -8.503    | -5.3981   | -4.3687    | -3.2749  | -4.0697  |

### Modelo 3

$$Y_{(ij)} \sim \text{Wiener}(\alpha_{(ij)}, \beta_{(ij)}, \tau_{(ij)}, \delta_{(p_{ij}, o_{ij})})$$

Table 3: Media de los parámetros modelados

| Modelo 3  |           |           |           |          |          |
|-----------|-----------|-----------|-----------|----------|----------|
| Parámetro | A         | B         | C         | D        | E        |
| $\alpha$  | 0.88881   | 0.96849   | 0.97048   | 0.95588  | 0.9952   |
| $\beta$   | 0.45816   | 0.50246   | 0.49924   | 0.5284   | 0.53707  |
| $\tau$    | 0.0021119 | 0.0061697 | 0.0010137 | 0.016375 | 0.015178 |
| $b_0$     | 10.486    | 6.9346    | 4.8691    | 1.8444   | 2.3766   |
| $b_1$     | -10.635   | -7.3161   | -2.0708   | -0.16917 | 1.3034   |
| $b_2$     | -5.3338   | -6.343    | -6.2612   | 0.36938  | -5.0773  |
| $b_3$     | -3.264    | -0.47105  | -3.4009   | -2.7109  | -2.5218  |
| $b_4$     | -3.2369   | 0.94515   | 2.0646    | -3.6848  | 1.1816   |

### Modelo 4

$$Y_{(ij)} \sim \text{Wiener}(\alpha_{(ij)}, \beta_{(ij)}, \tau_{(ij)}, \delta_{(p_{ij})})$$

Table 4: Media de los parámetros modelados

| Modelo 4  |           |           |           |          |          |
|-----------|-----------|-----------|-----------|----------|----------|
| Parámetro | A         | B         | C         | D        | E        |
| $\alpha$  | 0.83153   | 0.93395   | 0.94481   | 0.93427  | 0.96406  |
| $\beta$   | 0.46895   | 0.50739   | 0.49798   | 0.52976  | 0.53892  |
| $\tau$    | 0.0029797 | 0.0070839 | 0.0011369 | 0.017322 | 0.015736 |
| $b_0$     | 2.5307    | 2.1188    | 0.92531   | 0.34464  | -0.76552 |
| $b_1$     | 1.7668    | -4.2162   | 2.2647    | -0.33574 | 5.1825   |
| $b_3$     | -8.9003   | 0.55325   | -5.3849   | -0.91345 | -4.9913  |

### Modelo 5

$$Y_{(ij)} \sim \text{Wiener}(\alpha_{(ij)}, \beta_{(ij)}, \tau_{(ij)}, \delta_{(o_{ij})})$$

Table 5: Media de los parámetros modelados

| Modelo 5  |           |           |           |           |          |
|-----------|-----------|-----------|-----------|-----------|----------|
| Parámetro | A         | B         | C         | D         | E        |
| $\alpha$  | 0.76539   | 0.90863   | 0.92702   | 0.9406    | 0.9907   |
| $\beta$   | 0.46397   | 0.51086   | 0.49501   | 0.53197   | 0.53686  |
| $\tau$    | 0.0044623 | 0.0079759 | 0.0012098 | 0.017156  | 0.015165 |
| $b_0$     | 0.3456    | 0.21885   | 1.6944    | -0.073462 | 2.3286   |
| $b_2$     | 0.70658   | -0.45442  | -5.3103   | 2.6437    | -6.063   |
| $b_4$     | -0.037301 | 0.27331   | 4.1839    | -4.4153   | 2.5567   |

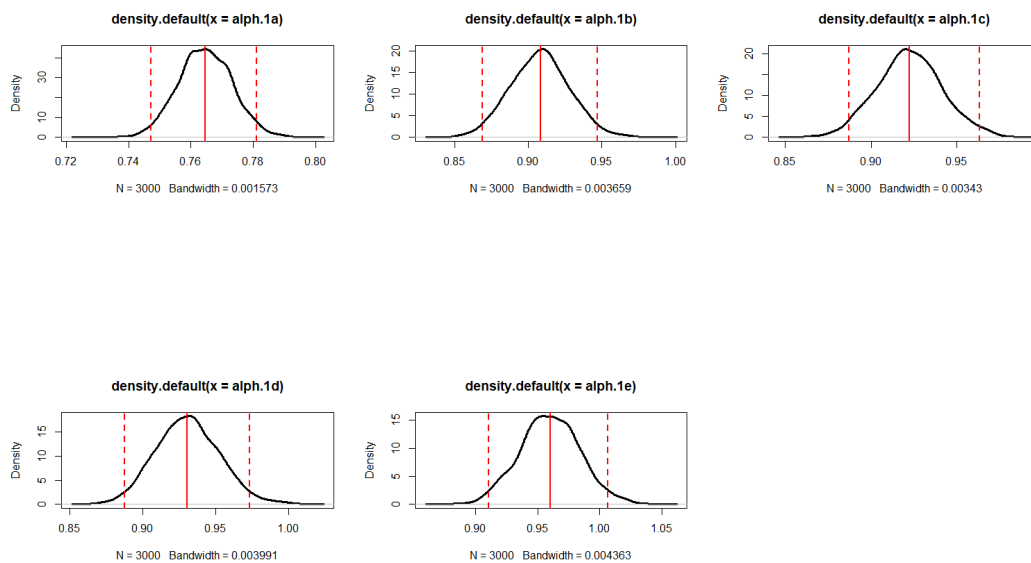
# Resultados - Análisis de parámetros comparados por modelo

## Parámetro $\alpha$

### Modelo 1

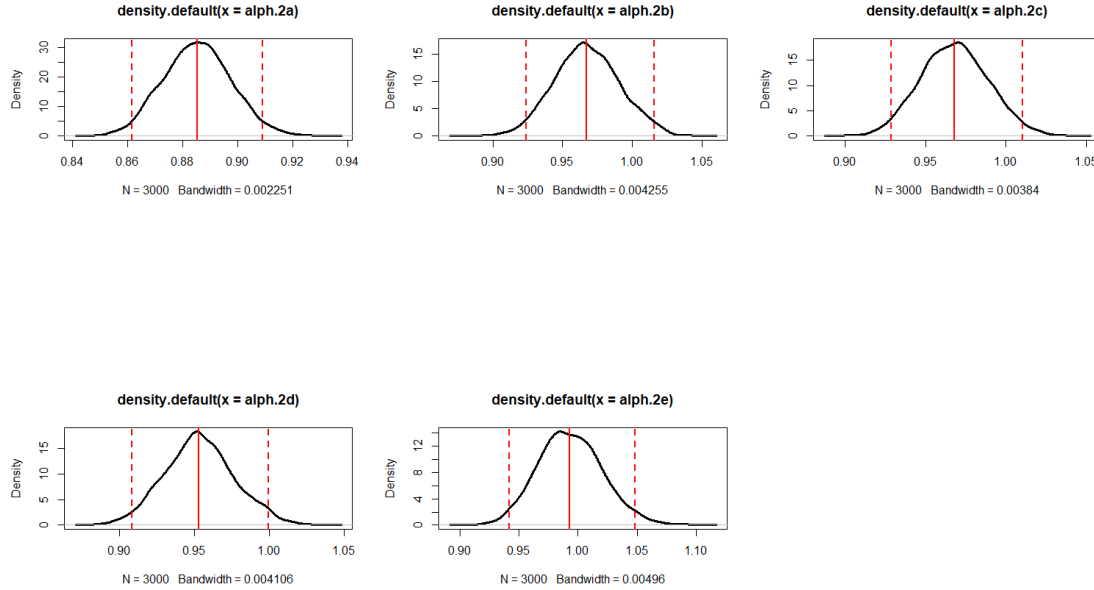
Mayores niveles de  $\alpha$  (*trade-off Speed-accuracy*) está relacionado con un mayor nivel de accuracy que tuvieron los participantes. Si comparamos todos los  $\alpha$ , como se muestran en la figura 1, desde la figura 1 “alpha.1a” (sin ambigüedad) hasta la “alpha.1.e” (mayor nivel de ambigüedad); es posible apreciar que entre menor grado de ambigüedad, más bajo es el nivel de  $\alpha$ , es decir, menor accuracy, el que iría aumentando a medida que el contraste incrementa.

Esto podría deberse a que en “alpha.1a”, la velocidad podría verse disminuida ya que los participantes pueden ver con claridad (no hay ambigüedad) la barra de la ruleta con los colores, y así, “tomarse más tiempo” para tomar una decisión; lo que no ocurriría en alpha.1.e, ya que la barra tiene una máscara que oculta el verdadero color de esas áreas.



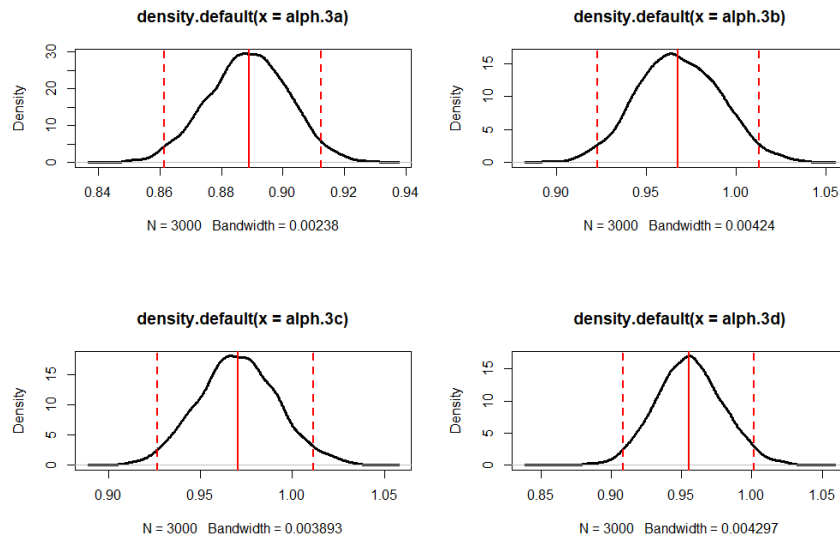
## Modelo 2

Con respecto a los  $\alpha$ , a nivel general, al igual que en el modelo 1, cuando no se presenta el contraste o ambigüedad, el trade-off entre el tiempo y accuracy es menor, a que cuando hay ambigüedad.



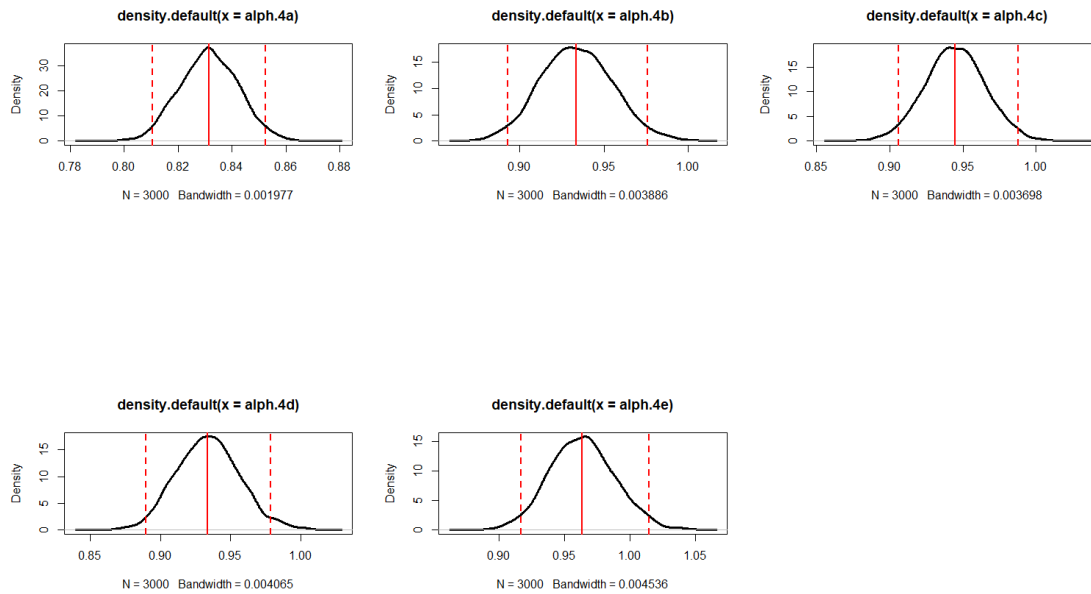
## Modelo 3

En relación a los  $\alpha$ , al igual que en los modelos anteriores (1 y 2), cuando no hay presencia de ambigüedad (contraste que oscurece la barra), el trade-off entre el tiempo y accuracy es menor a cuando si hay presencia de ambigüedad.



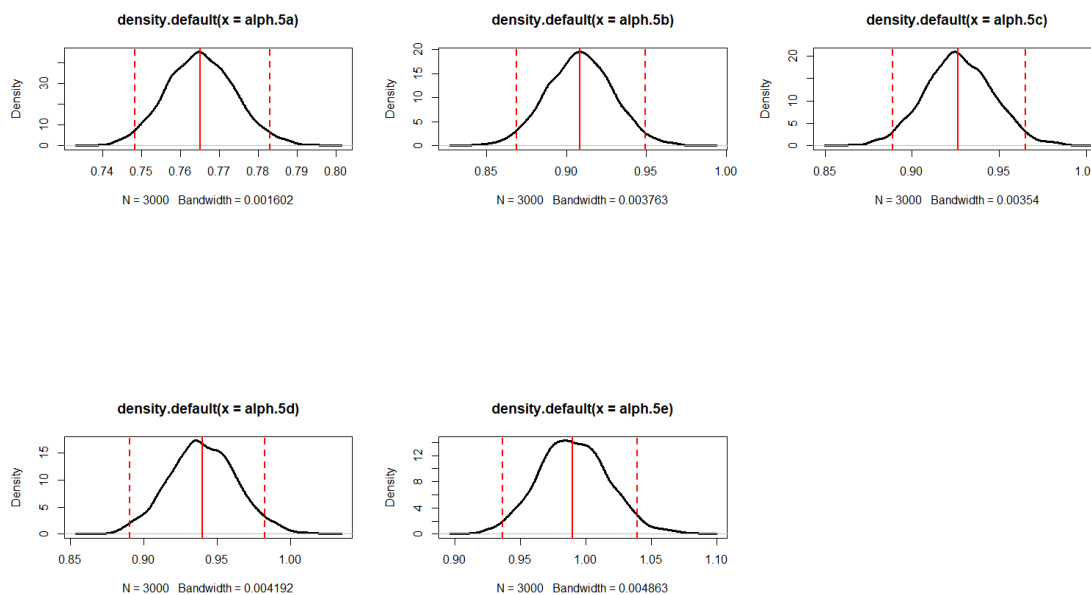
## Modelo 4

En relación a los  $\alpha$  del modelo 3, al igual que en los modelos anteriores (1,2 y 3), cuando no hay presencia de ambigüedad (contraste que oscurece la barra), el trade-off entre el tiempo y accuracy, es menor a cuando si hay ambigüedad.



## Modelo 5

Con respecto a los  $\alpha$ , a nivel general, al igual que en los modelos 1,2,3 y 4, cuando no se presenta el contraste o ambigüedad, el trade-off entre el tiempo y accuracy es menor a que cuando hay ambigüedad, o el contraste es ambiguo.

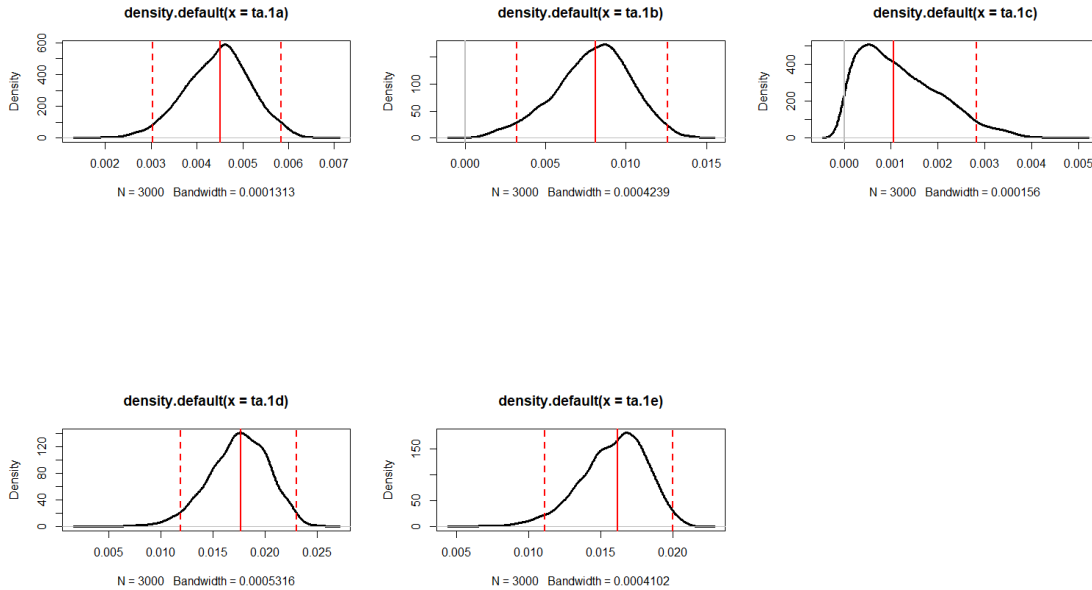




## Parámetro $\tau$

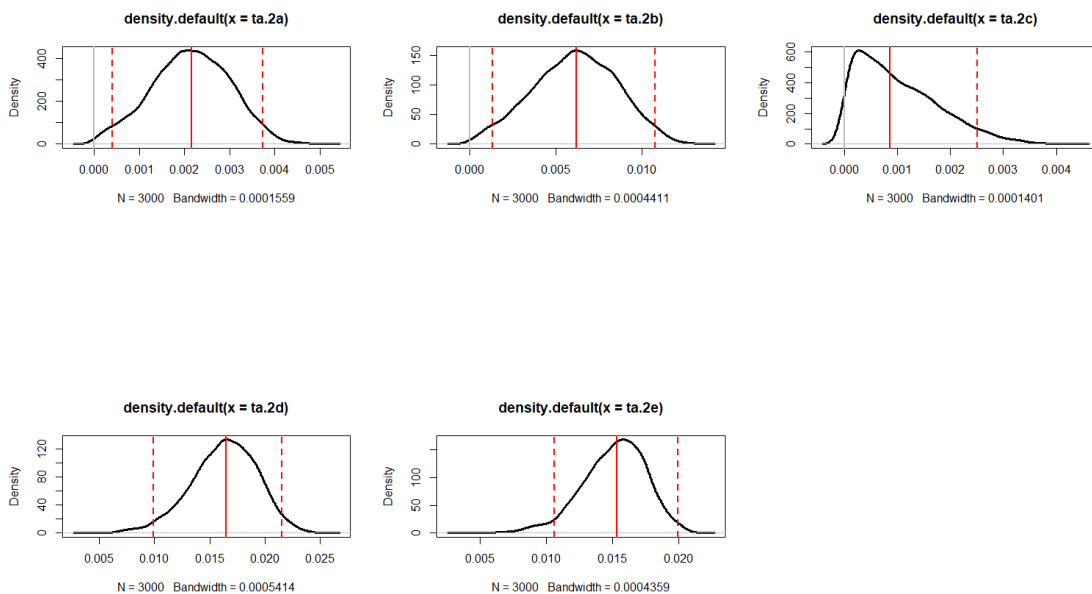
### Modelo 1

Finalmente, con respecto a los “tau” o tiempos de no-decisión (tiempo de respuestas motoras, tiempo de codificación, entre otros), los participantes tienen mayores tiempos de no-decisión cuando hay un mayor contraste (que cuando no lo hay).



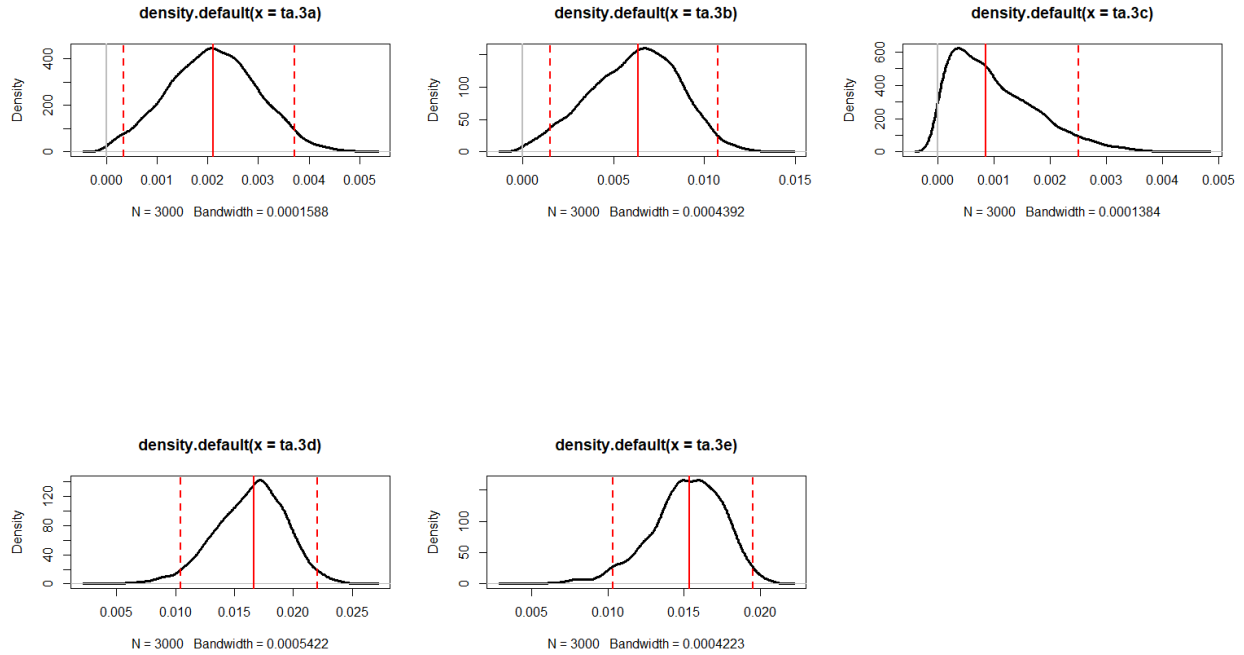
### Modelo 2

Con respecto a los “tau” o tiempos de no-decisión, los participantes tienen mayores tiempos de no-decisión cuando hay un mayor contraste, al igual que en el tau del modelo 1.



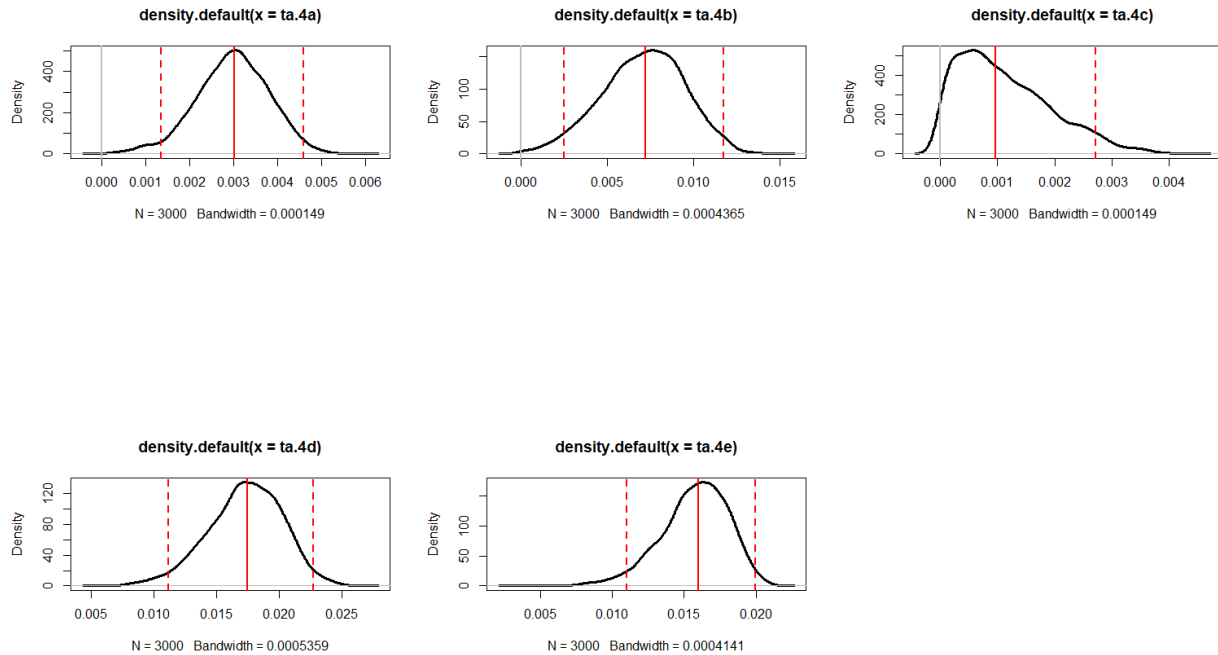
### Modelo 3

Con respecto a los “tau” o tiempos de no-decisión, los participantes tienen mayores tiempos de no-decisión cuando hay un mayor contraste.



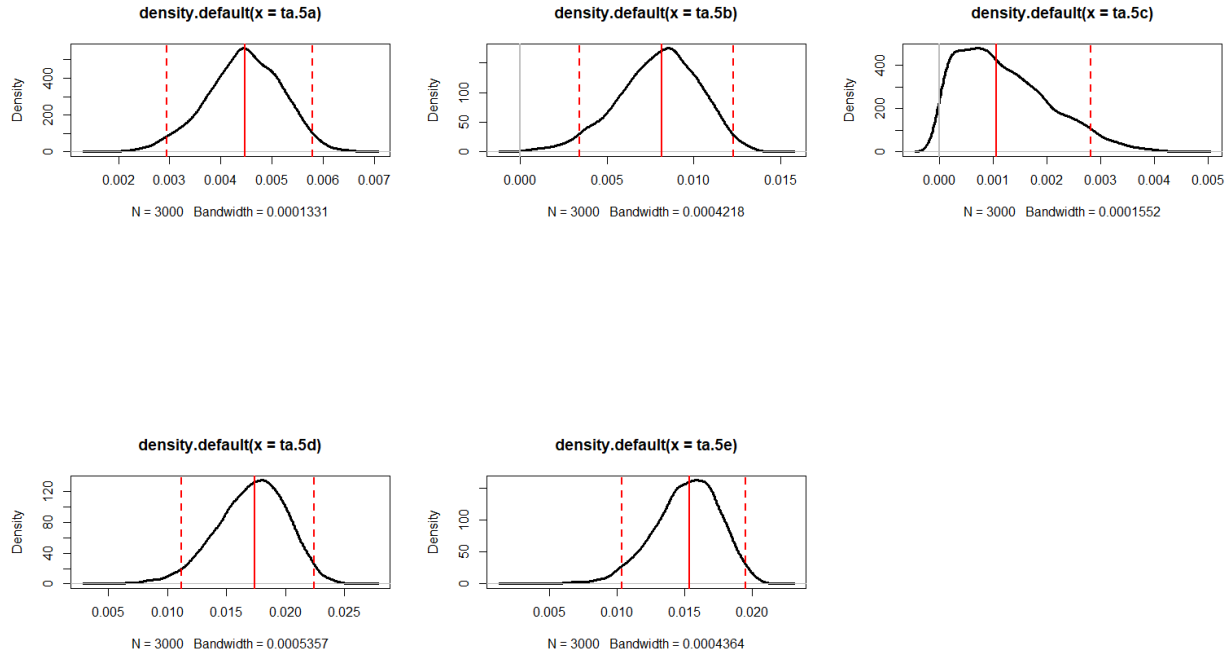
### Modelo 4

Con respecto a los “tau” o tiempos de no-decisión, los participantes tienen mayores tiempos de no-decisión cuando hay un mayor contraste, patrón que se repite en los modelos anteriores.



## Modelo 5

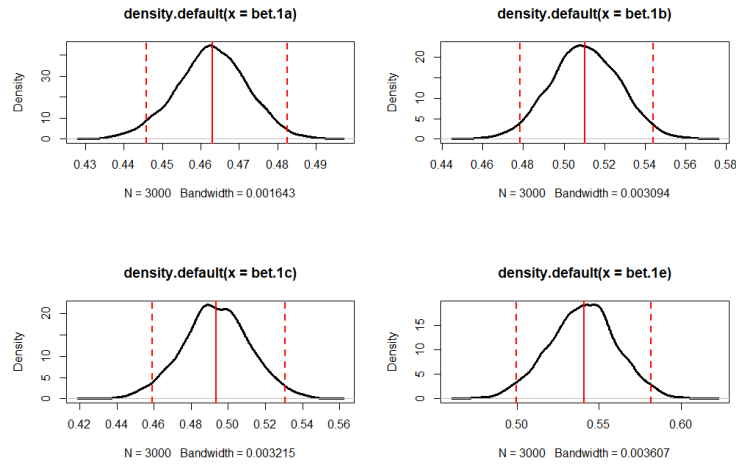
Con respecto a los “tau” o tiempos de no-decisión, los participantes tienen mayores tiempos de no-decisión cuando hay un mayor contraste.



## Parámetro $\beta$

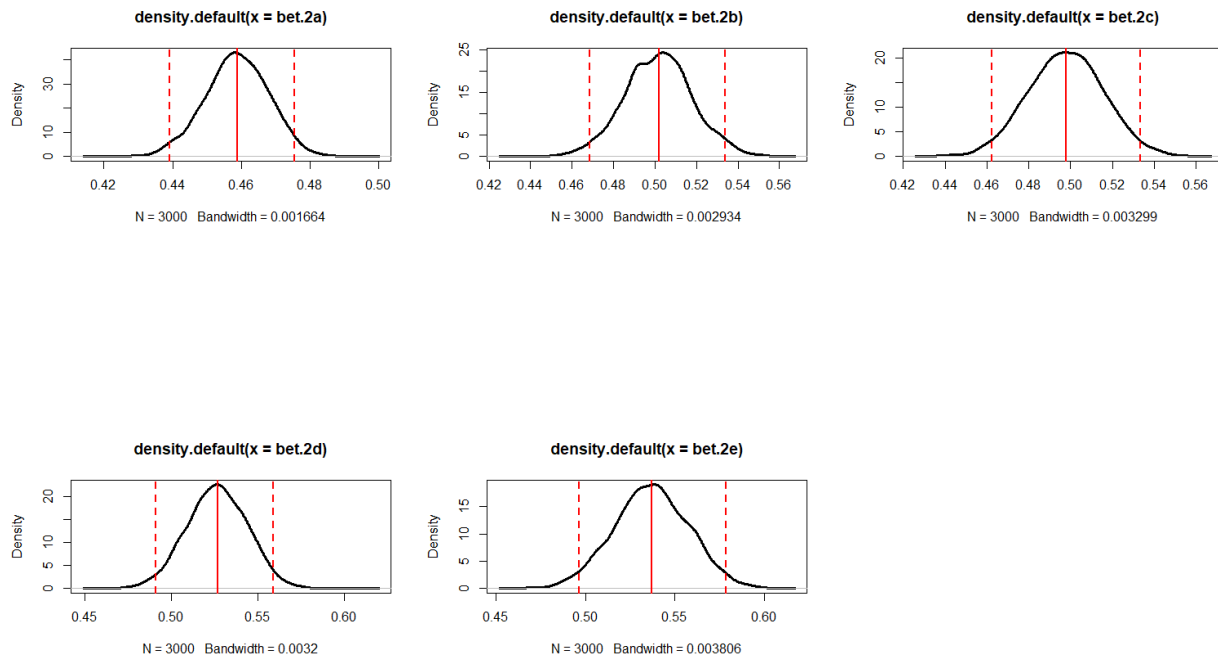
### Modelo 1

Respecto a los  $\beta$ , cuando no se presenta ambigüedad, es posible apreciar que existe cierto margen de sesgo (*bias*) hacia presionar el botón de la ruleta izquierdo. Sin embargo, a medida que se va presentando ambigüedad, el sesgo deja de presentarse.



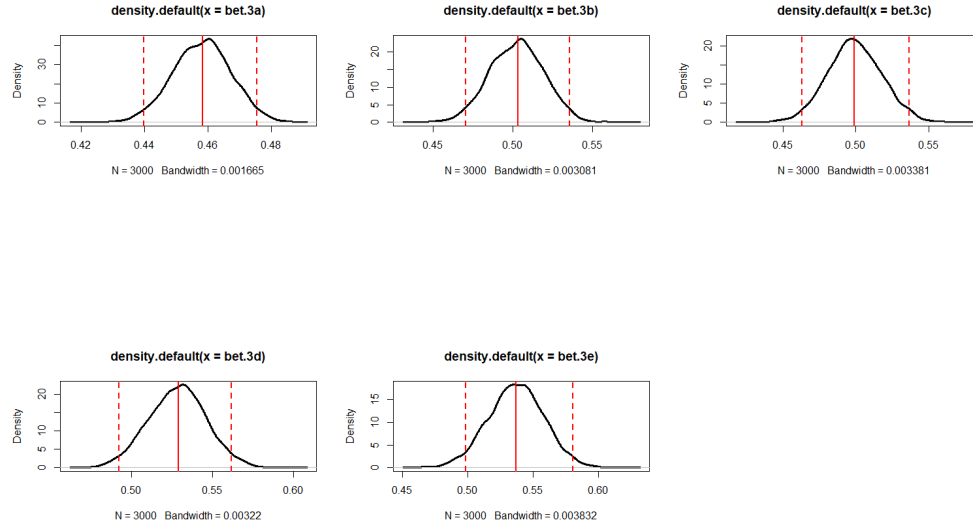
### Modelo 2

Respecto al  $\beta$  o *bias inicial*, cuando no hay ambigüedad, nuevamente (al igual que el Beta del modelo 1), existe un sesgo hacia presionar la barra izquierda, versus la derecha. Fenómeno que tendería a desaparecer a medida que el contraste de la barra y/o ambigüedad que se va presentando aumenta.



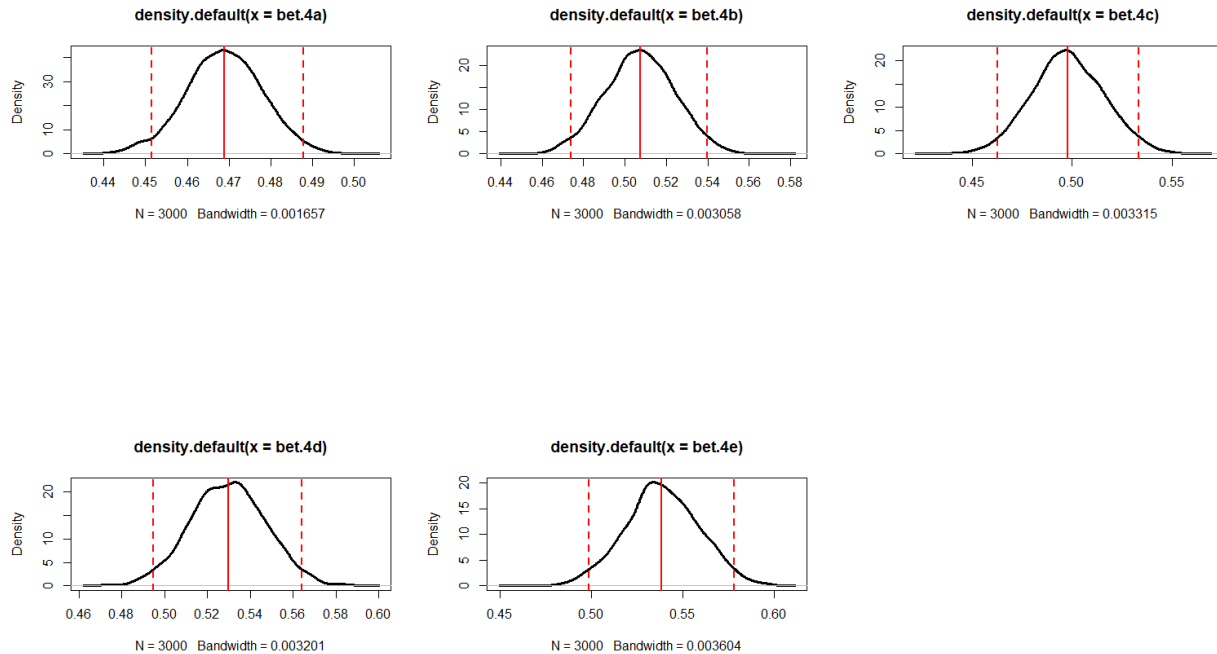
### Modelo 3

Respecto al  $\beta$ , cuando no hay ambigüedad, nuevamente, existe un sesgo hacia presionar la barra izquierda, versus la derecha. Fenómeno que tendería a desaparecer a medida que el contraste de la barra y/o ambigüedad que va teniendo el participante, aumenta.



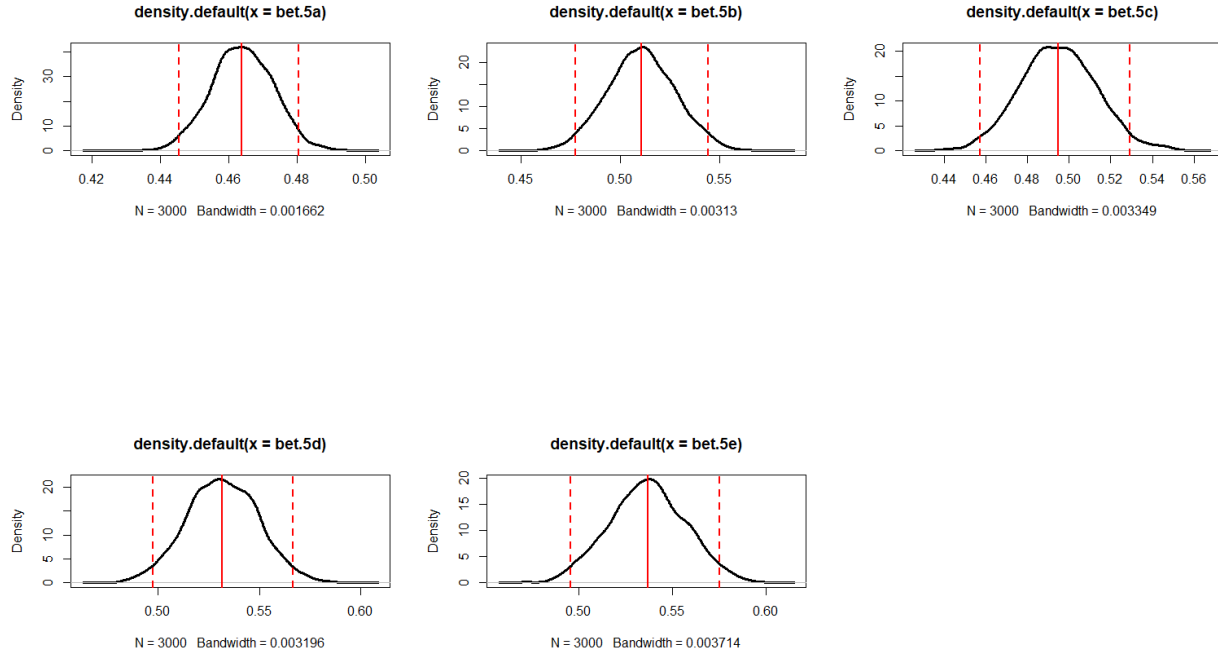
### Modelo 4

Respecto al  $\beta$  o bias inicial, cuando no hay ambigüedad, nuevamente, existe un sesgo hacia presionar la barra izquierda, versus la derecha. Fenómeno que tendería a desaparecer a medida que el contraste de la barra y/o ambigüedad que va teniendo el participante, aumenta.



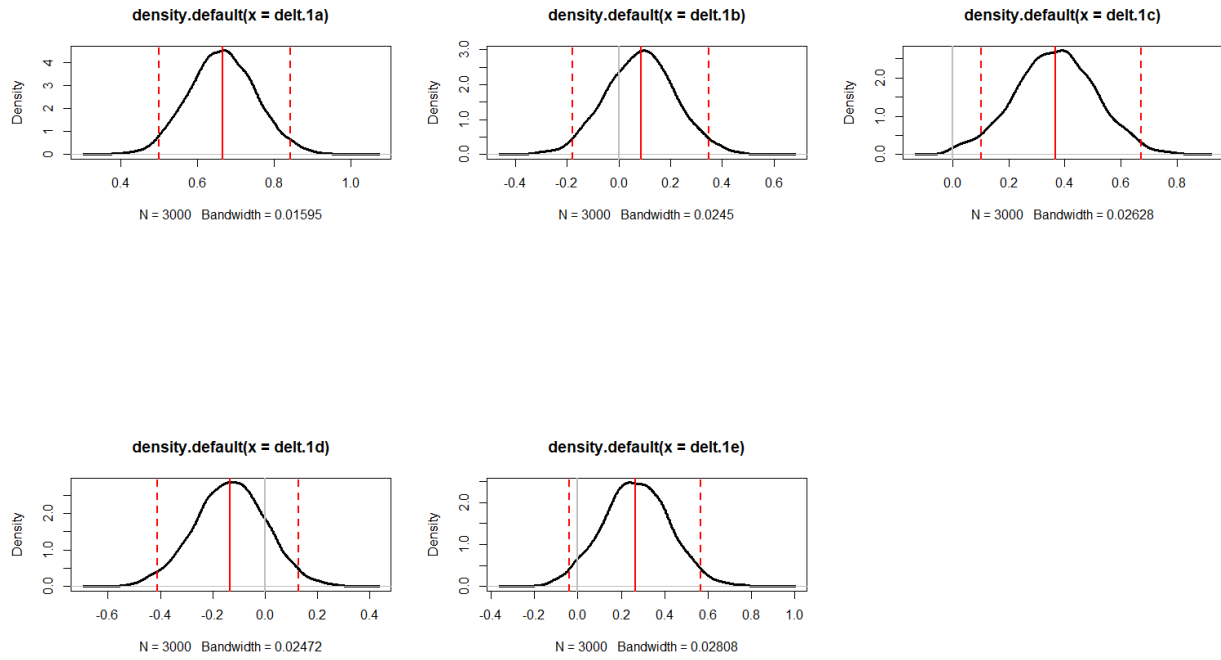
## Modelo 5

Respecto al  $\beta$  o bias inicial, cuando no hay ambigüedad, nuevamente, existe un sesgo hacia presionar la barra izquierda, versus la derecha. Fenómeno que tendería a desaparecer a medida que el contraste de la barra y/o ambigüedad que va teniendo el participante, aumenta.



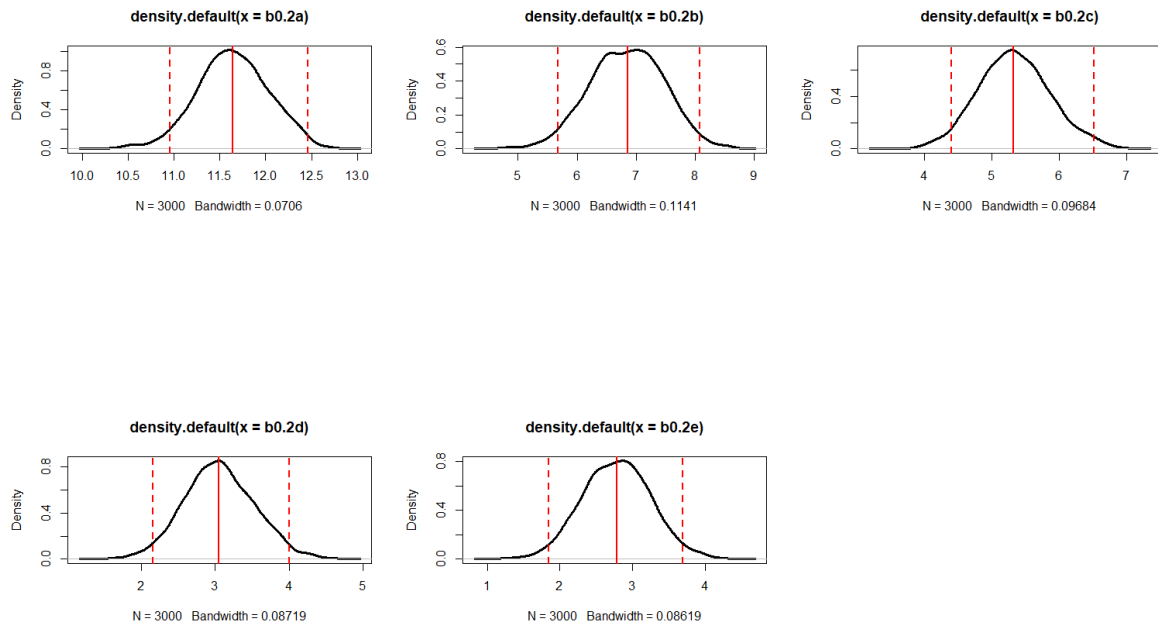
## Parámetro $\delta$

### Modelo 1

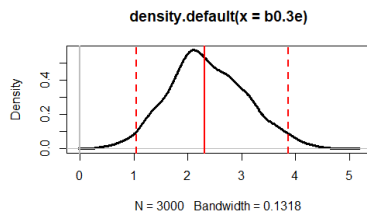
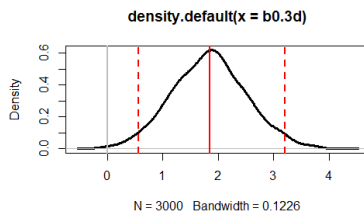
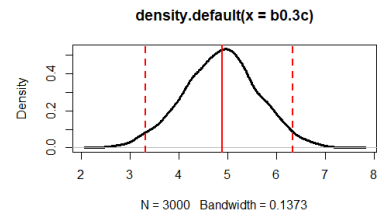
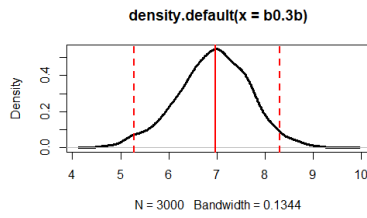
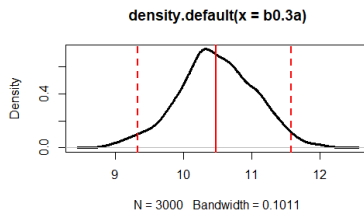


## Parámetro B0

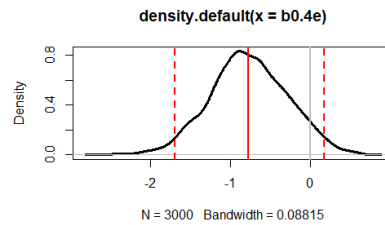
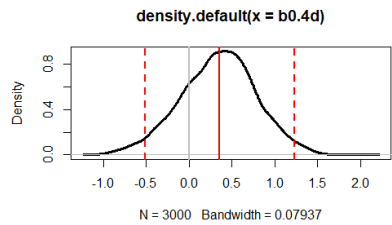
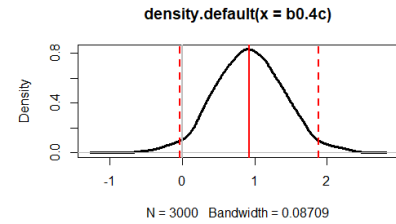
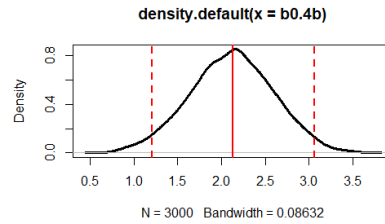
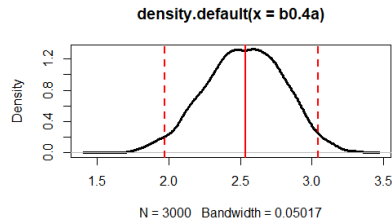
### Modelo 2



## Modelo 3

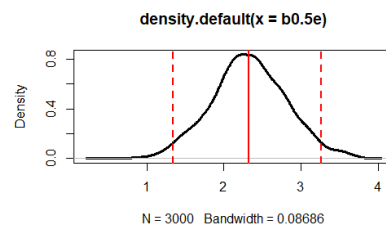
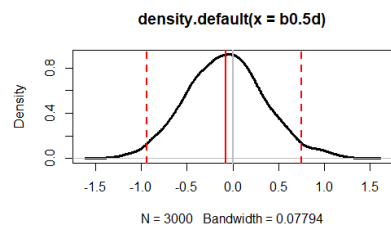
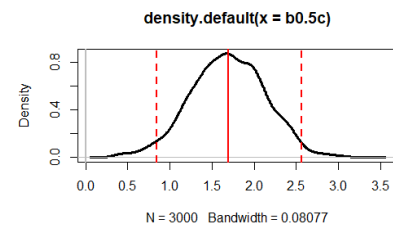
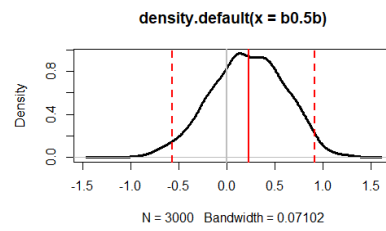
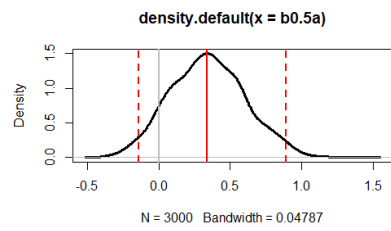


## Modelo 4





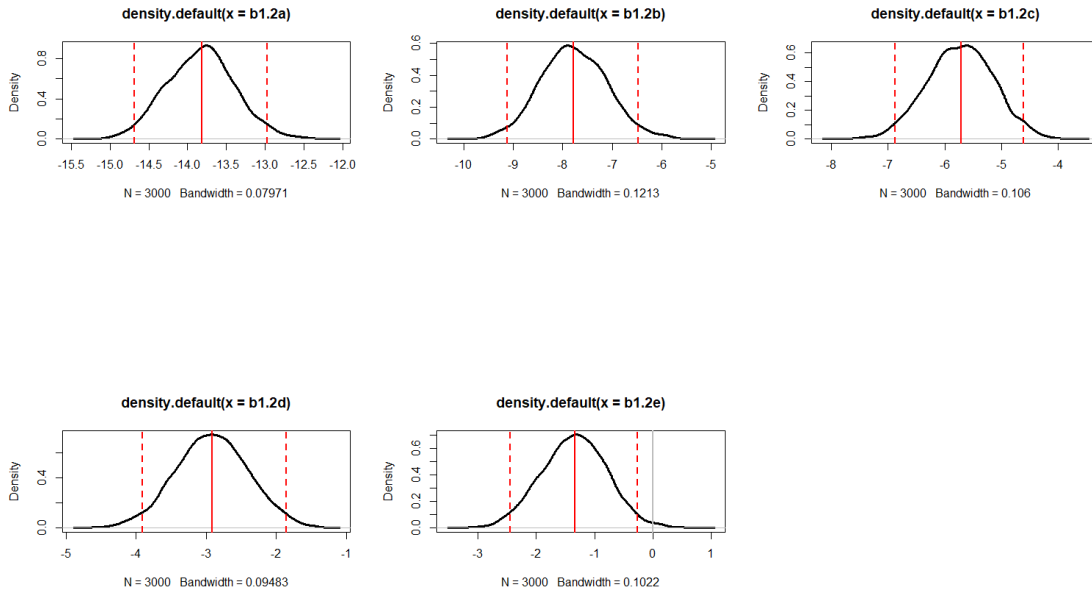
## Modelo 5



## Parámetro B1

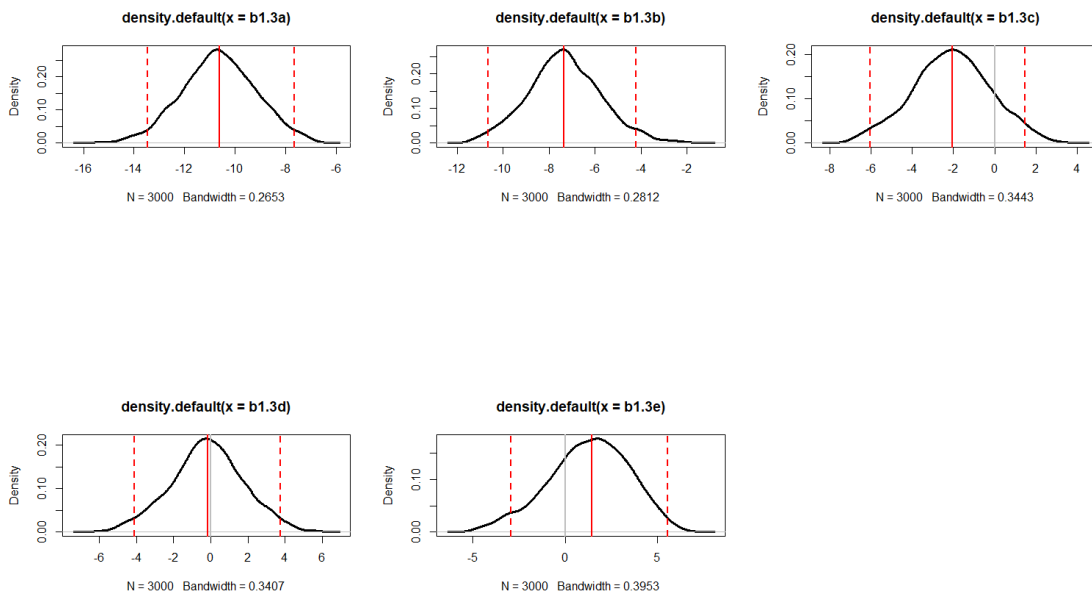
### Modelo 2

De acuerdo a los b1, tanto en condiciones de ambigüedad como de no ambigüedad, la probabilidad que acompaña al b1 si influye en el cálculo del drift.



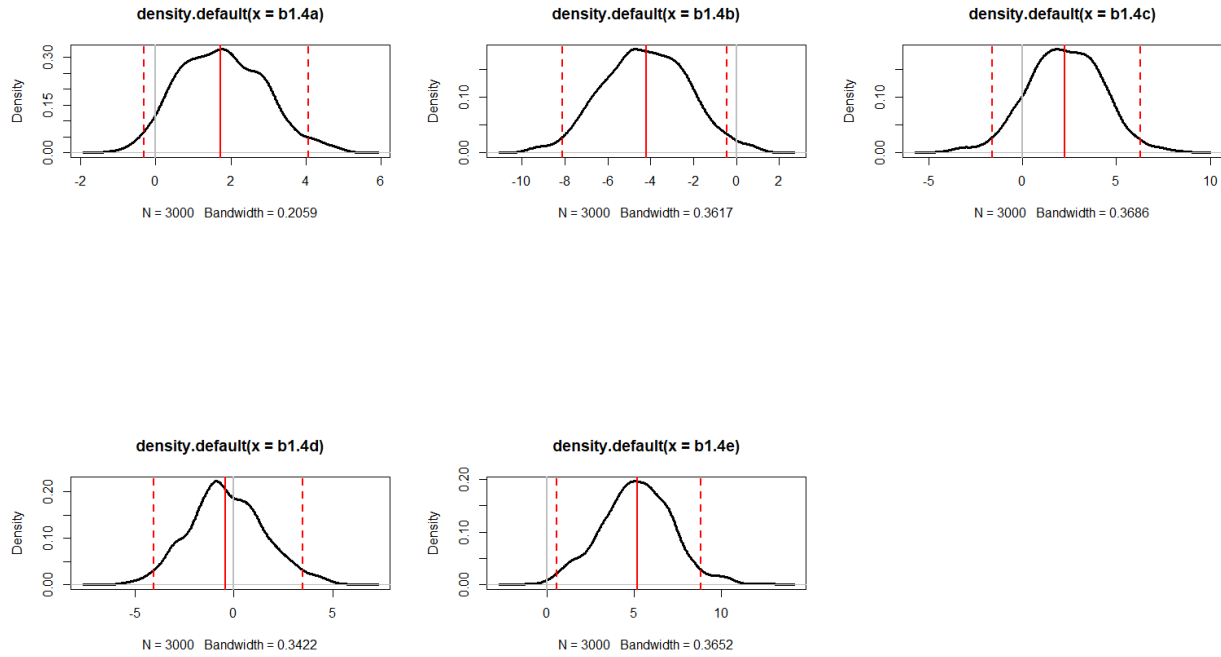
### Modelo 3

De acuerdo a los b1, a diferencia del modelo 1 y 2, ahora, a medida que se va presentando un aumento de contraste (o ambigüedad), la probabilidad que acompaña al b1 deja de influir en el cálculo del drift.



## Modelo 4

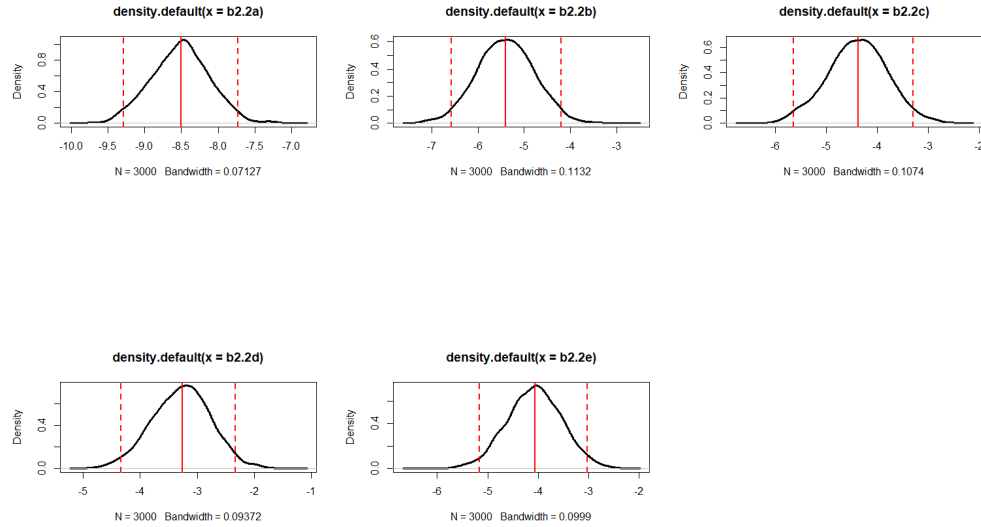
De acuerdo a los b1, tanto en condiciones de ambigüedad como de no ambigüedad, la probabilidad que acompaña al b1 no influye en el cálculo del drift, sin embargo, esto cambia según el contraste que se les presente a los participantes.



## Parámetro B2

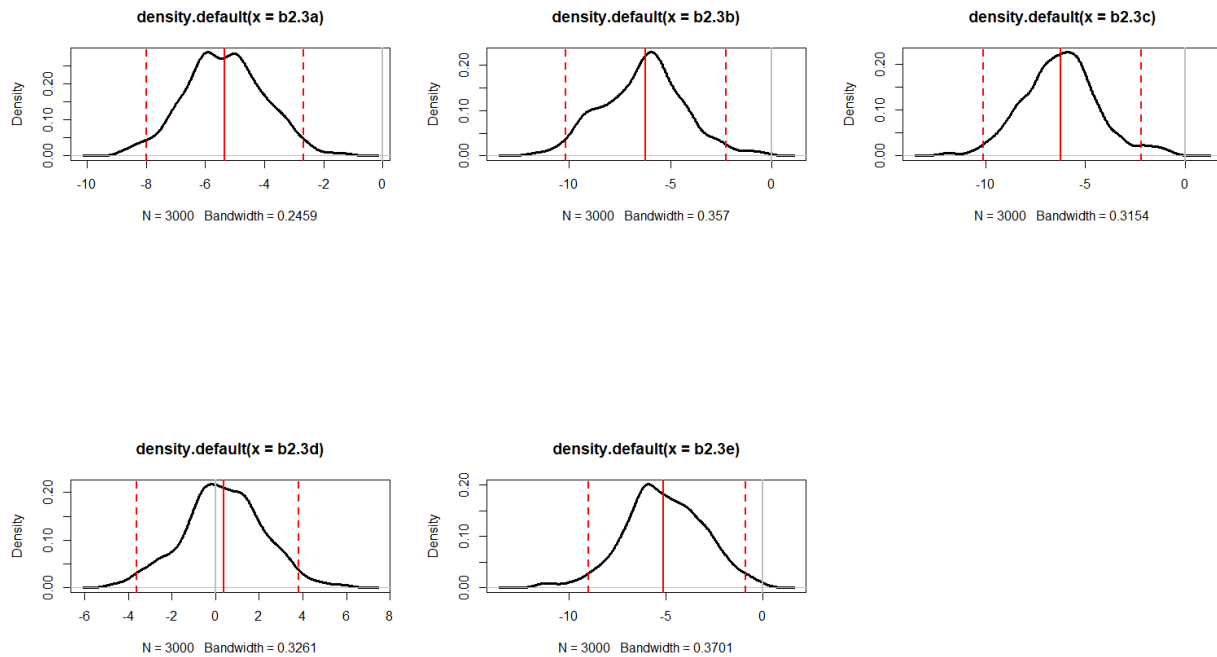
### Modelo 2

Finalmente, respecto a b2, tanto en condiciones de ambigüedad como de no ambigüedad, el pago que acompaña al b2 si influye en el cálculo del drift.



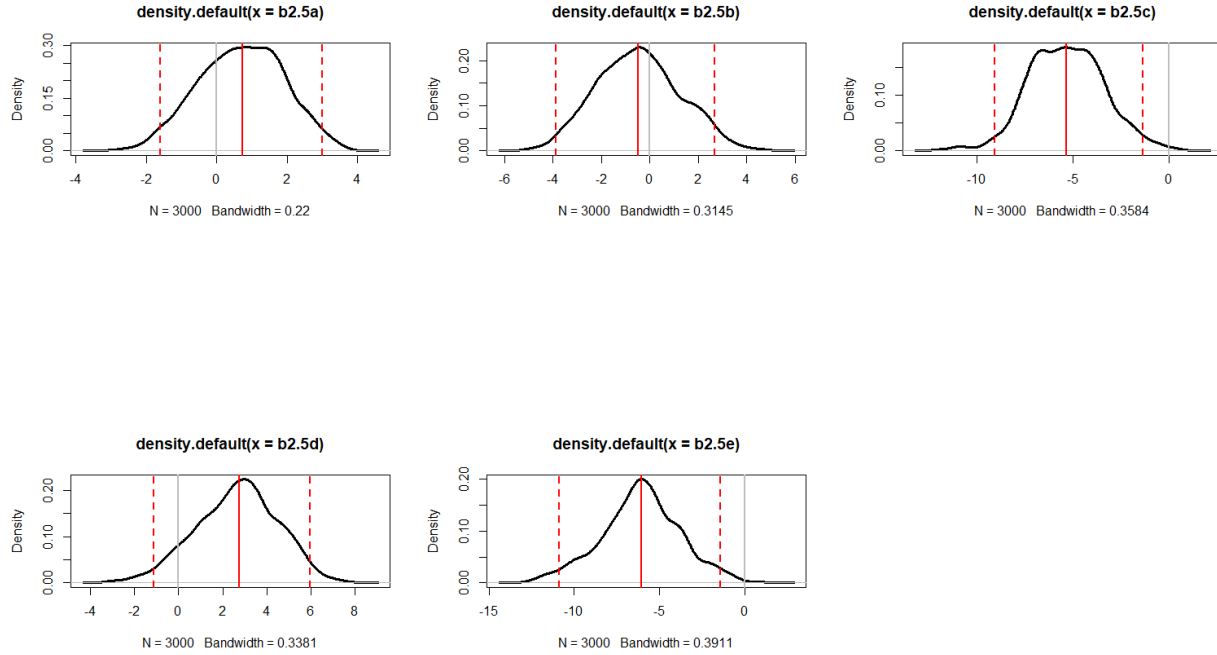
### Modelo 3

A nivel general el pago que acompaña a b2 si influye en el cálculo del drift, a excepción del b2.3d.



## modelo 5

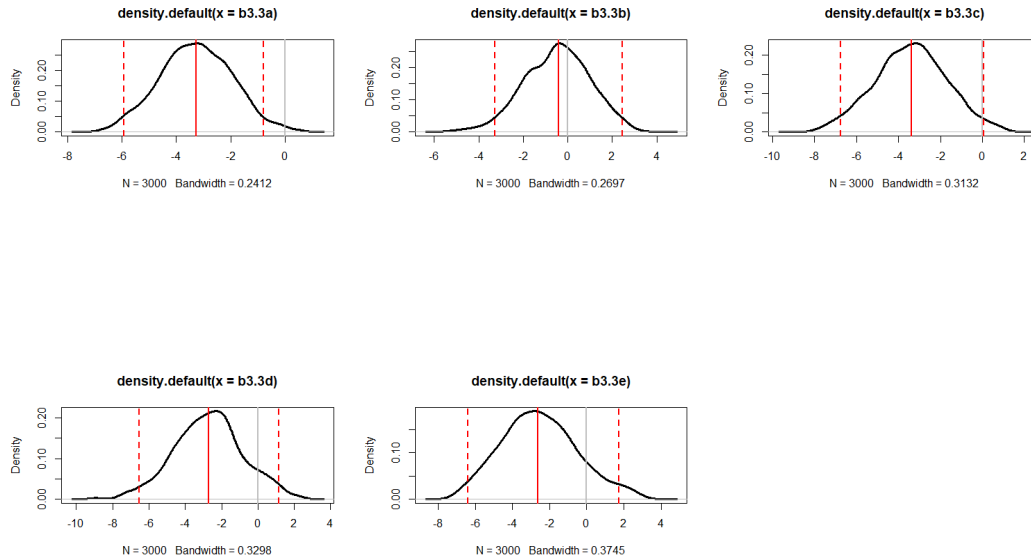
De acuerdo a los b2 (modelo no incluye ni b1 ni b3), a nivel general, el pago que acompaña a b2, influye y no influye en el cálculo del *drift*, dependiendo del nivel de ambigüedad (o no) que se presenta a los participantes. Por ejemplo en b2.5a, b2.5b, y b2.5d si influye, pero en b2.5c y b2.5e no influye.



## Parámetro B3

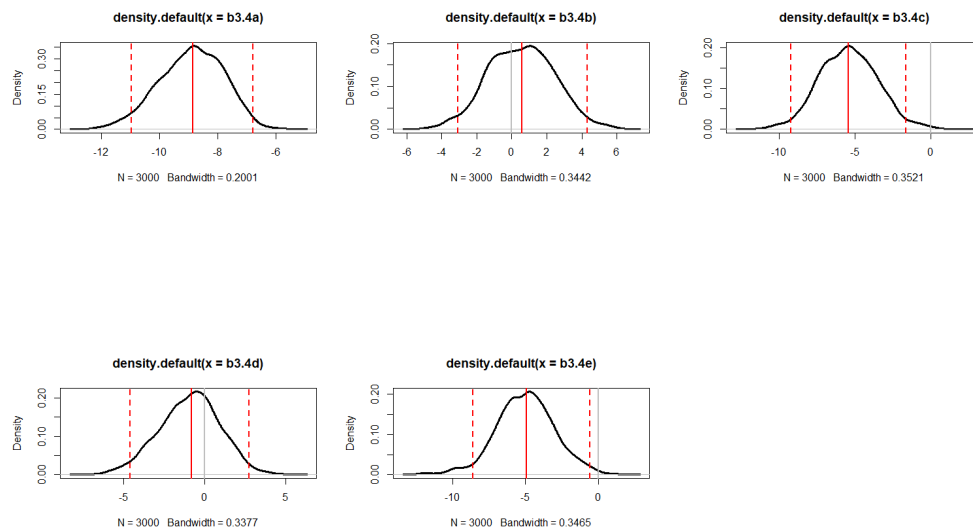
### Modelo 3

De acuerdo a los b3, que es una función cuadrática de la probabilidad; es posible apreciar que a medida que aumenta la ambigüedad la probabilidad de los b3 no influye en el cálculo drift.



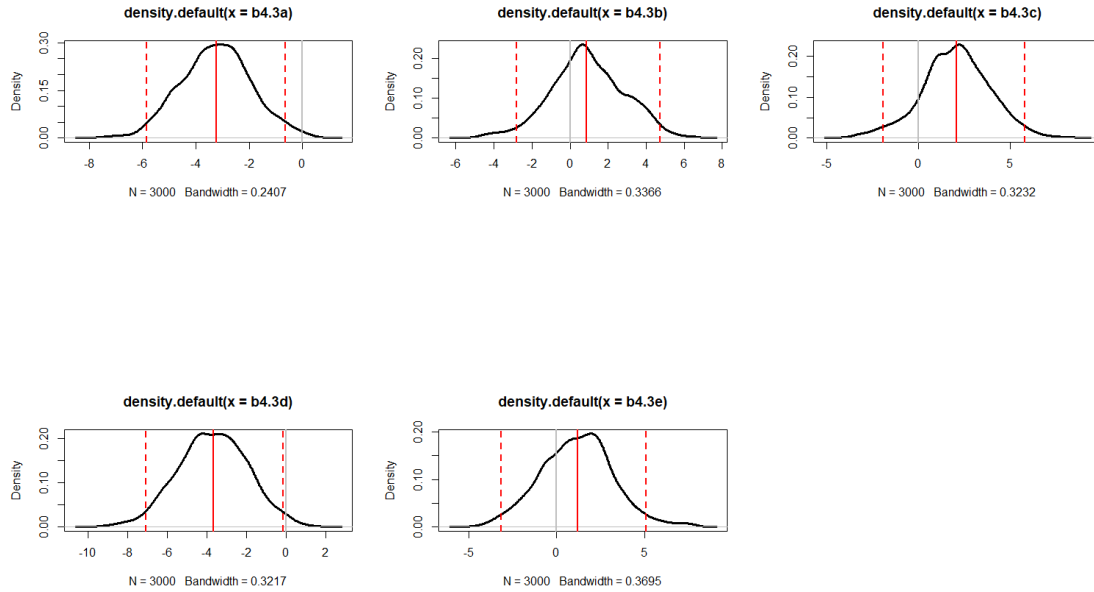
### Modelo 4

Finalmente, de acuerdo a los b3 (el modelo no incluye b2 y b4), sólo en condiciones de ambigüedad, la probabilidad que acompaña al b3 no influye en el cálculo del drift, sin embargo, esto cambia según el contraste en el que estén los participantes.



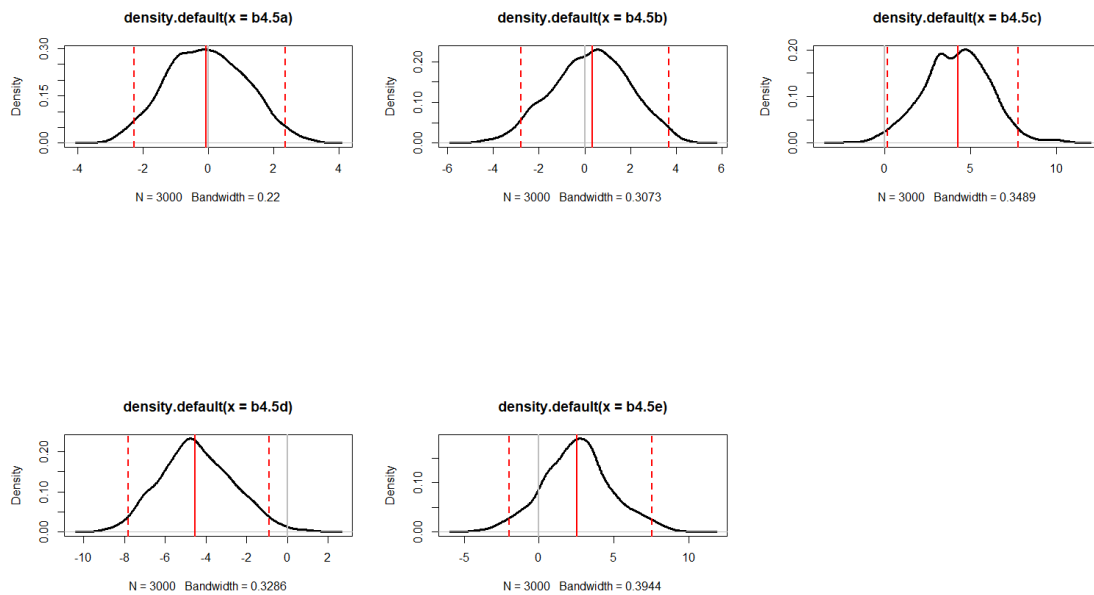
## Parámetro B4

**Modelo 3** Finalmente, con respecto a los b4, que es una función cuadrática del pago, es posible apreciar que a medida que aumenta la ambigüedad (al igual que el caso anterior) el pago no influye en el drift.



## Modelo 5

Finalmente, con respecto a los b4, de manera similar al b2, hay casos en que el pago influye en el drift, y casos en que no, independientemente del nivel de ambigüedad.



## Ajuste de modelo DIC y LOO

Como queda en evidencia en la tabla 6, el DIC indica que en el experimento sin máscara (A), el modelo que mas se ajusta a los datos es aquel que incorpora la probabilidad y el premio en una relación cuadrática. Este hallazgo se repite para todos los niveles de ambigüedad, exceptuando la máscara al 0.6 (E) y y al 0.3 (B). No obstante, esta diferencia es marginal.

| Table 6: Resultados DIC |       |      |      |     |     |
|-------------------------|-------|------|------|-----|-----|
| Mascaras                |       |      |      |     |     |
| Modelo                  | A     | B    | C    | D   | E   |
| 1                       | -1367 | -31  | -76  | 25  | -20 |
| 2                       | -2367 | -168 | -176 | -16 | -71 |
| 3                       | -2390 | -166 | -176 | -19 | -69 |
| 4                       | -1977 | -92  | -126 | 21  | -22 |
| 5                       | -1369 | -29  | -82  | 7   | -65 |

Los hallazgos resultantes del DIC son concordantes al revisar el Leave-One-Out (LOO), dado que el modelo 3 sigue siendo el de mejor ajuste en función a la muestra trabajada, con una diferencia marginal con el modelo 2 (el cual muestra un ajuste ligeramente mejor en las máscaras al 0.3 y 0.6).

| Table 7: Resultados LOO |      |    |    |     |    |
|-------------------------|------|----|----|-----|----|
| Mascaras                |      |    |    |     |    |
| Modelo                  | A    | B  | C  | D   | E  |
| 1                       | 683  | 16 | 36 | -13 | 8  |
| 2                       | 1188 | 84 | 86 | 8   | 34 |
| 3                       | 1195 | 82 | 86 | 9   | 33 |
| 4                       | 988  | 46 | 62 | -11 | 9  |
| 5                       | 684  | 14 | 40 | -4  | 31 |

Los hallazgos resultantes del DIC son concordantes al revisar el Leave-One-Out (LOO), dado que el modelo 3 sigue siendo el de mejor ajuste en función a la muestra trabajada, con una diferencia marginal con el modelo 2 (el cual muestra un ajuste ligeramente mejor en las máscaras al 0.3 y 0.6).

Finalmente, como próximos desafíos, sería interesante ahondar en otras funciones para la relación de la probabilidad y el premio con el Drift.

## Referencias

Wabersich, D., & Vandekerckhove, J. (2014). Extending JAGS: A tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behavior Research Methods*, 46(1), 15–28. <https://doi.org/10.3758/s13428-013-0369-3>