

## Parte III

# Criterios de evaluación

## Capítulo 8

# *L* y *S* variables

*Habiendo descrito en el capítulo anterior la forma en la que llevaremos a cabo la estrategia de exploración secuencial de diseños —mediante factoriales y teniendo en cuenta 3 variables de valor fijo: L, S y w— llevada a cabo para 15 cuadros de puntos distintos, aquí continuaremos configurando los datos hacia el mismo objetivo de nuestro problema. Nos proponemos en este capítulo repetir la misma evaluación para los 15 cuadros de puntos del capítulo anterior, pero aquí en los 25 casos que resultan de combinar 5 niveles para la variable L y otros 5 niveles para la variable S, manteniendo fija la tercera variable, w. Esta variable es la que nos permite partir el estudio determinando un mismo primer centro de experimentación en todos los casos. De este modo, ampliaremos el alcance de nuestro estudio y evaluaremos en cada caso el comportamiento de los criterios de medida de calidad de ajuste.*

### 8.1. Introducción

Para comenzar a entender mejor el comportamiento de las 3 variables de estudio que hemos propuesto —*w*, *L* y *S*— cuando toman distintos niveles, en este capítulo comenzaremos por establecer 5 niveles para *L* y para *S*, manteniendo fijo el nivel que utilizamos para la tercera en el capítulo anterior, es decir,  $w = 5\%$ . Esta última situación nos llevará a considerar tres consideraciones básicas:

- a. Considerar los  $5 \times 5 = 25$  casos de estudio que esto traerá como consecuencia.
- b. Que cada uno de estos 25 casos de estudio se basará en la misma mecánica de exploración que seguimos en el capítulo anterior, utilizando también el presupuesto fijo equivalente a 15 puntos de diseño ó 1500 observaciones “éxito-fracaso”:
  - todas las exploraciones secuenciales hacia el máximo partirán de un mismo centro de experimentación inicial, y

- cada uno de los 25 casos de estudio contendrá la información proveniente de respectivos 15 cuadros de puntos, igualmente a como hicimos en el capítulo anterior con un caso particular.
- c. Que cada uno de estos 25 casos de estudio seguiremos la misma orientación que en el capítulo anterior: buscar qué niveles de  $L$  y de  $S$  son los que mejores valores nos dan para los dos criterios de evaluación propuestos como medidas de validación de los ajustes conseguidos.

A lo largo de este capítulo, plantearemos un esquema de estudio para esta propuesta con  $L$  y  $S$  variables —de allí, el título de este capítulo— y preparar el terreno para el capítulo siguiente, en donde daremos 5 niveles para la última variable,  $w$ , teniendo al mismo tiempo en cuenta los 5 niveles que dimos a  $L$  y a  $S$  en este capítulo.

## 8.2. Definición de la matriz $LS$

Una forma de agrupar en un mismo “objeto” los 5 niveles que tomarán las variables  $L$  y  $S$  es mediante la definición de una simple matriz de 25 filas por 2 columnas, en la que las filas representen cada uno de los 25 casos de estudio que consideraremos, es decir, los 25 pares  $(L, S)$  distintos que surgen de dar niveles a ambas variables de estudio. A esta matriz la denominaremos simplemente “matriz  $LS$ ”.

### 8.2.1. Sobre la elección de niveles para $L$ y $S$

Como habíamos comentado en capítulos anteriores, la falta de información acerca de puntos reales de funcionamiento del proceso que estudiamos —el cual simulamos computacionalmente— nos ha conducido a realizar algunas suposiciones y simplificaciones en el estudio que nos ocupa. Una de ellas —o, más bien, *dos* de ellas— tienen que ver con los valores que elegimos tanto para la variable  $L$ , como para  $S$ . Estas dos cantidades representan: (a) el valor que tendrán los lados de los diseños que utilizemos —de forma cuadrada— y (b) el módulo del salto entre dos centros de diseño sucesivos, respectivamente. Como también indicábamos, los cosenos directores de estos saltos se determinan tomando parte de la estrategia clásica de  $MSR$ , es decir, a partir de las estimaciones de los coeficientes estimados del modelo que agrupa los 15 puntos de diseño de cada cuadro de puntos.

El punto de partida para elegir qué niveles daremos para ambas variables no resulta trivial, si bien observamos que su dificultad puede reducirse relativamente si partimos

de un principio conceptual<sup>1</sup> que también proponemos, que se basa en que *teniendo ya definidos los niveles —o pares de valores ( $L, S$ )— para estas variables, el propósito de nuestro trabajo es el de proponer una forma objetiva de evaluar cuál de ellos es el que se acerca mejor al objetivo que nos hemos planteado, que se basa en maximizar los dos criterios que propusimos. Es así como el mayor énfasis lo pondremos en diseñar estas estrategias de evaluación suponiendo un conjunto de niveles, sin tener certeza que los mismos resultan los más convenientes entre todos los conjuntos posibles de niveles.*

Como último punto con el cual intentamos una justificación de la elección de las filas de la matriz  $\mathbf{LS}$ , nos apoyamos también en el viejo y no menos efectivo pensamiento que todo experimentador conviene que tenga a la hora de enfrentarse con nuevos problemas. Así, diremos que el proceso de *aprendizaje de los datos*<sup>2</sup> también está presente en nuestra concepción de la matriz  $\mathbf{LS}$ , en tanto y en cuanto que una vez que encontremos los “mejores valores” para las variables de estudio, no hay nada que impida redefinir una dirección de estudio que pueda buscar los “niveles óptimos” de ambas<sup>3</sup>, tarea que puede generalizarse sin demasiada dificultad mediante el uso de aplicaciones estadísticas potentes, como por ejemplo  $\mathbf{R}$ . De esta manera, ya podemos plantear una futura línea de investigación a partir de lo que podamos sacar en limpio de nuestro trabajo y de continuar aprendiendo cada vez mejor de los datos que nos hemos propuesto estudiar.

De esta manera, y conservando fijos tanto el vector  $\beta$  del capítulo anterior como el nivel  $w = 5\%$ , hemos elegido como punto de partida a la distancia  $d_{OM}$ , que separa el primer centro del máximo y que es medida sobre el plano de los factores. En el capítulo 6 hemos mostrado de qué forma se expresa la misma y una forma de calcularla mediante el uso de los programas de la serie  $\mathbf{BETAS}$ , que hemos escrito en  $\mathbf{R}$ .

---

<sup>1</sup>En el lenguaje de las organizaciones, este tipo de principio del cual hemos partido tiene la naturaleza de una *política*. Tomando el punto de vista de ROMERO MALETTI (1995), podemos decir que “una política es una necesidad de cuya satisfacción no se tiene certeza de conveniencia, y que se basa en criterios no tanto objetivos sino más bien en la experiencia”. A diferencia de las políticas, los *objetivos* no se basan en la experiencia sino en el *razonamiento objetivo*, lo cual lleva a que *se tenga certeza de la conveniencia de satisfacción de las necesidades*. En nuestro caso de definir los niveles para estas dos variables de estudio no tenemos criterios objetivos sobre los cuales basar los valores que propondremos, con lo que estableceremos éstos teniendo en cuenta aspectos conceptuales relacionados con un problema del que tenemos poca o ninguna información. Es así como hacemos una propuesta no exhaustiva para tomar estos niveles.

<sup>2</sup>Vid. ALUJA Y MORINEAU (1999).

<sup>3</sup>Podríamos conjeturar lo anterior diciendo que el mayor énfasis de nuestro trabajo lo hemos puesto en el *cómo llegar* más que en el *cuáles niveles*, aunque esta conjetura no es rígida en absoluto.

**Niveles para  $L$** 

Habiendo realizando unas pruebas previas<sup>4</sup> con  $\mathbf{R}$ , nos encontramos que si consideramos como valor central del vector de niveles para  $L$  al 100 % de la distancia  $d_{OM}$ , los ajustes conseguidos no nos aportaban información tan adecuada que si los comparamos con los obtenidos para fracciones menores, como por ejemplo, el 25 % de  $d_{OM}$ . Es por estos motivos que hemos elegido como niveles para  $L$  a un conjunto de valores centrados en la cantidad  $\frac{1}{4}d_{OM}$ , sin que ello signifique indicar que dicha selección de niveles es la óptima. Así, tendremos que los niveles que tomaremos para evaluar la variable “lado del diseño” serán los correspondientes al 60 %, 80 %, 100 %, 120 % y 140 % de dicha cantidad, es decir:

$$\text{Niveles de } L = \frac{1}{4} d_{OM} \times (0,6; 0,8; 1,0; 1,2; 1,4)'$$

**Niveles para  $S$** 

Del mismo modo como partimos de una solución de compromiso —pero con cierto fundamento— para establecer los niveles de  $L$ , razonamos análogamente para proponer un vector de niveles para el módulo del salto entre centros sucesivos,  $S$ . Nos ha parecido útil plantearnos una pregunta de partida: *¿qué pasaría si el salto lo tomáramos igual a la distancia  $d_{OM}$ ?* En este caso, y dependiendo de la calidad del ajuste conseguido, sería factible aproximarnos al máximo de un solo salto si es que los cosenos directores resultaron los adecuados. Sin embargo, si estos cosenos no fueron lo suficientemente informativos, también podríamos observar que nos podríamos desviar de la “dirección correcta” en una cantidad relativamente grande, de módulo  $d_{OM}$ .

Es por esto que nos pareció adecuado iniciar el estudio dando valores a la variable  $S$  que resulten de fracciones situadas entre el 20 % y el 100 % de la distancia  $d_{OM}$ , de modo que tendremos así los 5 niveles para el salto:

$$\text{Niveles de } S = d_{OM} \times (0,2; 0,4; 0,6; 0,8; 1,0)'$$

Para la superficie en cuestión y con la ayuda de  $\mathbf{R}$ , los niveles para ambas variables quedan definidos en los vectores columna  $\mathbf{LL}$  y  $\mathbf{SS}$ :

```
> LL
[1] 0.2064356 0.2752474 0.3440593 0.4128711 0.4816830
> SS
[1] 0.2752474 0.5504948 0.8257422 1.1009896 1.3762370
```

<sup>4</sup>La forma en como hemos escrito los programas de las series  $\mathbf{BETAS}$  o  $\mathbf{CRJER}$  (ver **Apéndice C**), permiten realizar pruebas sin dificultad alguna, variando solamente el valor dado a una variable.

**La matriz LS**

Finalmente podemos formar todos los casos combinados entre ambos vectores<sup>5</sup>, con los cuales armaremos la matriz **LS**:

```
> LS <- expand.grid(LL, SS)
> colnames(LS) <- c("Valores de L", "Valores de S")
```

```
> LS
  Valores de L Valores de S
1    0.2064356  0.2752474
2    0.2752474  0.2752474
3    0.3440593  0.2752474
4    0.4128711  0.2752474
5    0.4816830  0.2752474
6    0.2064356  0.5504948
7    0.2752474  0.5504948
8    0.3440593  0.5504948
9    0.4128711  0.5504948
10   0.4816830  0.5504948
11   0.2064356  0.8257422
12   0.2752474  0.8257422
13   0.3440593  0.8257422
14   0.4128711  0.8257422
15   0.4816830  0.8257422
16   0.2064356  1.1009896
17   0.2752474  1.1009896
18   0.3440593  1.1009896
19   0.4128711  1.1009896
20   0.4816830  1.1009896
21   0.2064356  1.3762370
22   0.2752474  1.3762370
23   0.3440593  1.3762370
24   0.4128711  1.3762370
25   0.4816830  1.3762370
```

---

<sup>5</sup>Nótese que el caso que estudiamos en los dos capítulos anteriores corresponden al formado por los dos niveles más bajos de ambas variables, que será la *fila número 1* de la matriz **LS**.

A partir de esta matriz  $\mathbf{LS}$ , iremos evaluando los resultados fila por fila, tanto para el criterio I como para el criterio II. Este agregado de variabilidad en ambas variables enriquecerá el alcance que venimos dando al estudio, del que las observaciones que podamos hacer resultarán más sólidas y útiles que si hubiésemos considerado un solo caso.

### 8.2.2. Gráficos de evaluación

El esquema que seguiremos para evaluar cada uno de los criterios de medida de calidad de ajuste lo llevaremos a cabo definiendo para cada caso de estudio 3 gráficos, contruidos de la misma manera para facilitar su comparación. Utilizaremos aquí, para abreviar, la notación  $T$  para referirnos a cualquiera de los criterios de evaluación, ya sea  $\det(\mathbf{I}_F)$  o bien  $\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$ . De este modo, consideraremos 3 gráficos básicos:

1. *Diagrama de rejilla*, en el que definiremos una “rejilla” cuyos nodos resulten de la representación en un gráfico de dos dimensiones ( $L$  y  $S$ ) de cada una de las filas de la matriz  $\mathbf{LS}$ , colocando en cada uno de ellos el valor calculado para el estadístico en cuestión que refiere el criterio correspondiente.
2. *Proyección sobre  $T(S)$* , que será la proyección de todos los 25 puntos de cada estadístico sobre el plano  $T(S)$ .
3. *Proyección sobre  $T(L)$* , que será la proyección de todos los 25 puntos de cada estadístico sobre el plano  $T(L)$ .

La figura 8.1 de la página siguiente servirá de esquema práctico para ilustrar esta idea de proyecciones que comentamos.

En todos los estadísticos, representamos los valores promedio de los 15 cuadros de puntos correspondientes a cada una de las 25 filas de la matriz  $\mathbf{LS}$ , que los hemos generado nuevamente desde el inicio. Las instrucciones de programa con las que hemos realizado los mismos se encuentran en los programas de la serie MEDCA, que pueden verse en el **Apéndice C**.

## 8.3. Criterio I: cantidad de información

El primer criterio que habíamos definido para comparación de estrategias era el del determinante de la  $MIF$  que, como habíamos indicado, en el caso binomial depende de la matriz de pesos  $\mathbf{W}$ , la que a su vez es función de la probabilidad de éxito estimada,  $\hat{\pi}_i$ . Ésta la podemos calcular como la proporción observada,  $\text{pr. obs}$ , o bien como la probabilidad calculada con el modelo ajustado,  $\hat{\pi} = \frac{\exp(\hat{\eta})}{1 + \exp(\hat{\eta})}$ , con  $\hat{\eta} = \hat{\beta}_0 + \mathbf{x}'\hat{\mathbf{b}} + \mathbf{x}'\hat{\mathbf{B}}\mathbf{x}$ .

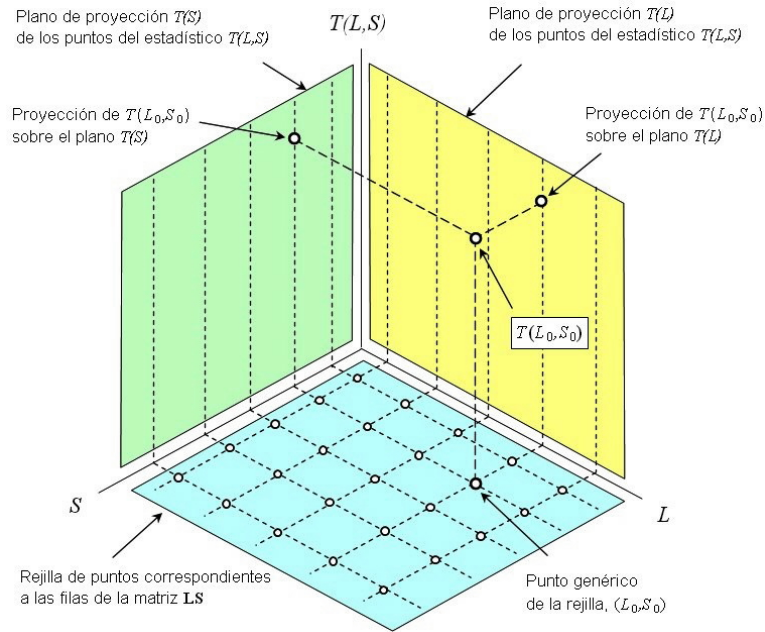


Figura 8.1: Mediante la notación  $T(L, S)$  se representan los estadísticos de comparación  $\det(\mathbf{I}_F)$  y  $\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$ , como así también las transformaciones que se definan a partir de ellos. En cada caso, se presentarán las figuras específicas.

En esta sección presentaremos unos “diagramas de rejillas” en donde definimos una cuadrícula para los 5 niveles de  $L$  y  $S$  considerados, de modo que cada uno de los nodos así determinados definen una fila de la matriz  $\mathbf{LS}$ . En cada uno de los nodos representaremos el valor promedio de los 15 cuadros de puntos correspondientes a la misma superficie de respuesta teórica y a los valores de  $L$  y  $S$  correspondientes, como así también el valor fijo  $w = 5\%$ .

Utilizando la notación  $T$ , nos referiremos de manera genérica a los estadísticos o medidas de evaluación que surgen de cada uno de los dos criterios propuestos. Así, tendremos:

- Para el criterio I, habrá 2 medidas de evaluación:  $\det(MIF)$  y  $\log[\det(MIF)]$ , cada una de ellas calculadas de acuerdo a los dos puntos de vista que muestran las ecuaciones (7.1) y (7.2).
- Para el criterio II, tendremos 4 medidas:  $\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$ ;  $\text{logit} [\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})]$ ;  $\text{asin} [\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})]$  y  $\text{fyt} [\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})]$ .

En las secciones siguientes, evaluaremos cada una de estas medidas, cuyo contexto general puede verse en la figura 8.2.



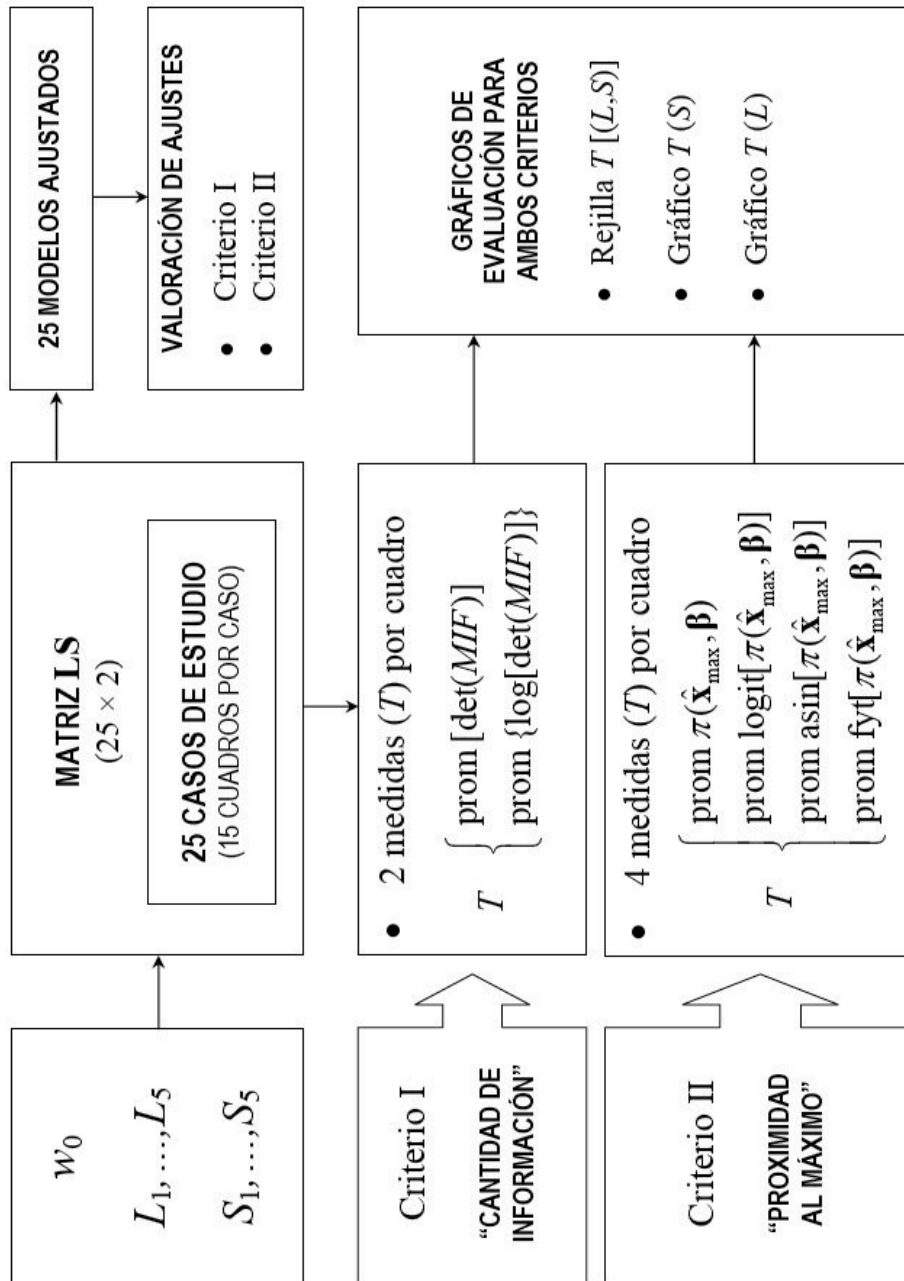


Figura 8.2: Situación correspondiente a los 25 casos de estudio que se generan al considerar 5 niveles para las variables  $L$  y  $S$ , manteniendo  $w = 5\%$ .

### 8.3.1. MIF: determinante y su logaritmo

Estimación utilizando  $\hat{\pi} = \text{pr. obs}$

El criterio de los determinantes de la MIF evaluado en cada uno de los niveles de la matriz  $\mathbf{LS}$  lo representamos las figuras 8.3 y siguientes, de las que podemos ver que el máximo valor del estadístico se observa para los niveles más altos de ambas variables. Los valores de cada uno de estos puntos, para cada nivel de  $L$  y  $S$ , quedan resumidos en la siguiente matriz:

> MM.01

	S1	S2	S3	S4	S5
L1	2.207294	1154.865	9874.518	171958.8	1560168
L2	17.403592	5665.180	97688.538	820938.6	5449033
L3	171.721348	10510.454	273058.447	3402761.5	45807775
L4	949.768482	23982.027	629976.713	16613593.5	140792521
L5	3216.861341	69662.030	1569881.409	31743922.8	164539317

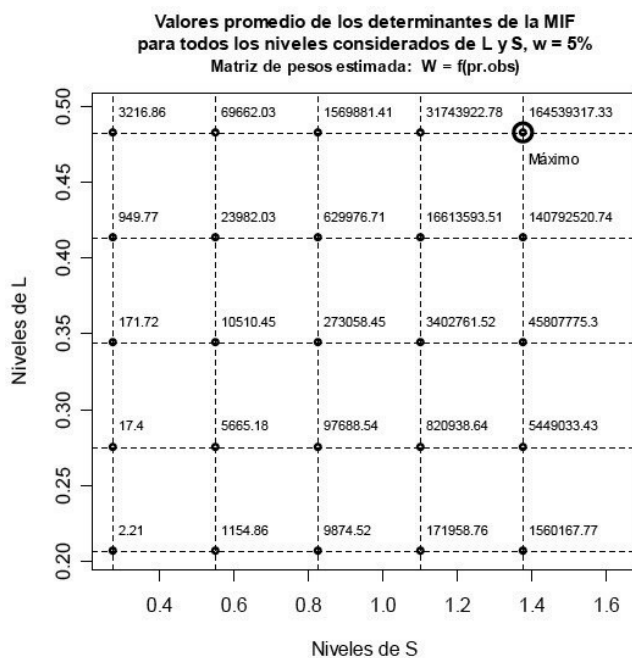
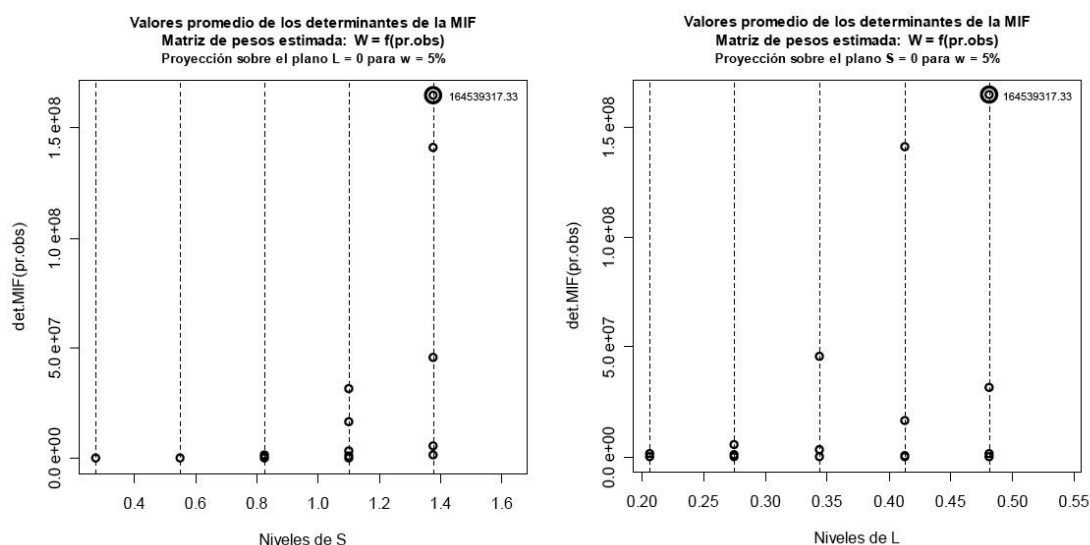


Figura 8.3: Rejilla de determinantes para todos los casos de  $L$  y  $S$ , estimando  $\pi$  con  $\text{pr. obs}$ .



Podemos ver en estos dos últimos gráficos que los mayores valores se obtienen para los niveles más altos de  $L$  y de  $S$ , si bien la variabilidad parece aumentar también en el mismo sentido en que aumenta el estadístico.

El mismo esquema de presentación de resultados para todas las filas de la matriz  $LS$  del punto anterior lo repetimos aquí para el logaritmo de los determinantes de la  $MIF$ , calculando la matriz de pesos de la misma manera. En la siguiente tabla presentamos sus valores para cada una de las filas de dicha matriz:

```
> MM.02
```

	S1	S2	S3	S4	S5
L1	0.609784	6.409485	8.789135	11.69656	13.68587
L2	2.832240	8.366841	10.771686	13.27071	14.90613
L3	4.920858	8.926051	11.966742	14.64284	17.23475
L4	6.710628	9.895784	12.858426	16.40311	18.31340
L5	7.836736	10.901976	13.928891	17.08853	17.95727

Las representaciones gráficas de los mismos las podemos ver en las figuras 8.4 y siguientes, en las que el máximo sigue prácticamente el comportamiento que en el caso anterior, salvo que el nivel correspondiente en  $L$  no es el más alto sino el inmediato anterior:

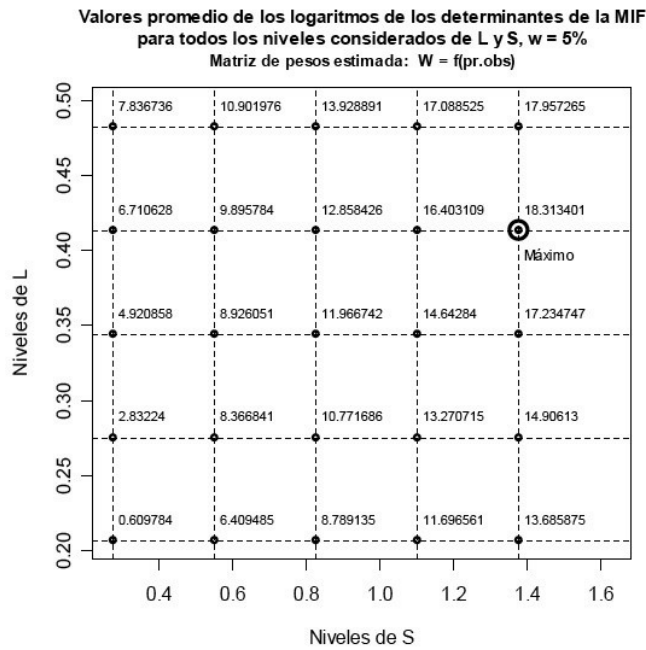
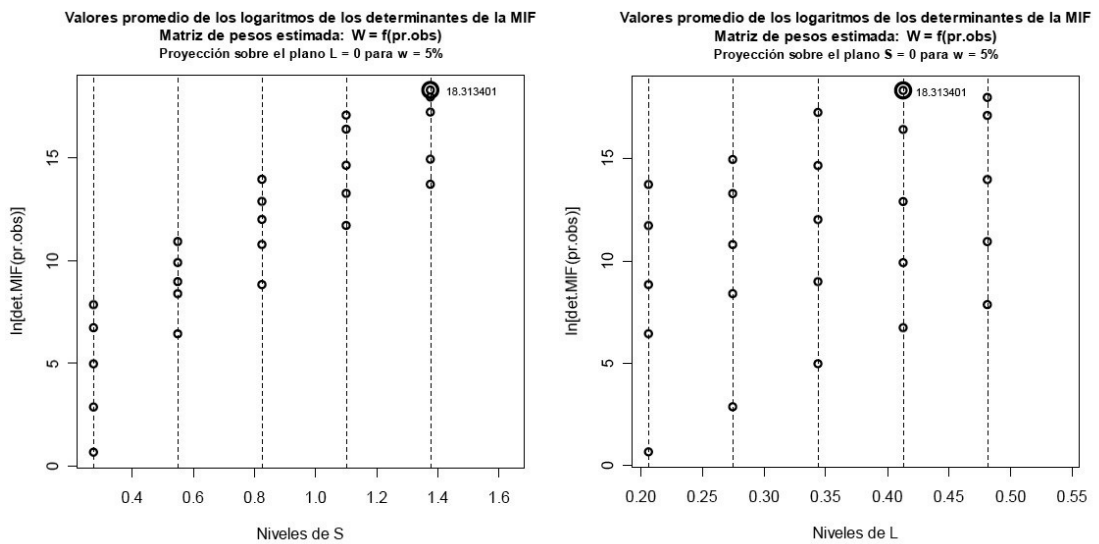


Figura 8.4: Rejilla de los logaritmos de los determinantes para todos los casos de  $L$  y  $S$ , estimando  $\pi$  con  $\text{pr.obs}$ .



Vemos en los dos últimos gráficos que la variabilidad de las observaciones del estadístico se mantiene relativamente estable a medida que aumentan los niveles de ambas variables, ya que es una de las características típicas de la función logaritmo.

### Estimación utilizando $\hat{\pi} = \text{pi.hat}$

La otra forma que indicamos que podíamos utilizar para calcular la probabilidad de éxito estimada en la matriz de “pesos” de la *MIF* era la de definirla a partir de la ecuación (7.2), valores que expresamos en la siguiente tabla:

```
> MM.03
```

	S1	S2	S3	S4	S5
L1	2.545291	959.2997	4419.829	1782.293	970777.8
L2	27.363937	5335.4845	48740.991	2895.362	9959669.4
L3	212.634377	5098.8209	27212.557	304590.155	20487868.7
L4	1076.207967	5380.6407	18415.072	17465932.644	89967586.8
L5	4449.719278	16721.2002	269690.235	10810273.785	107457520.8

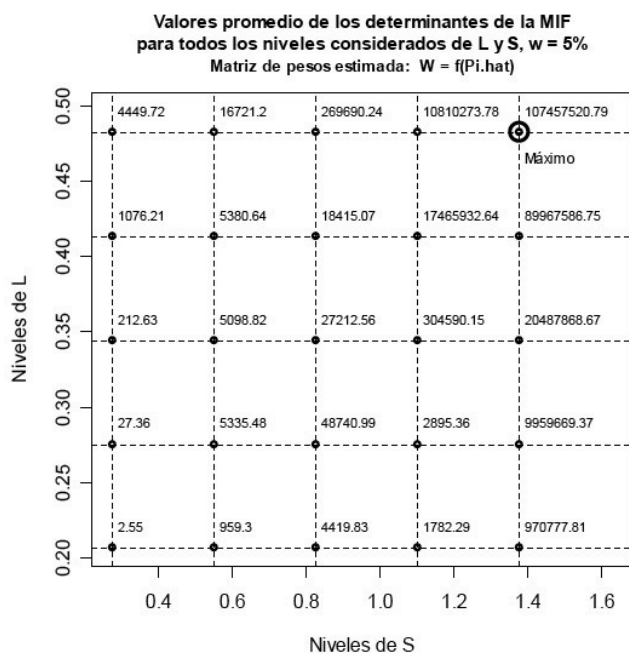
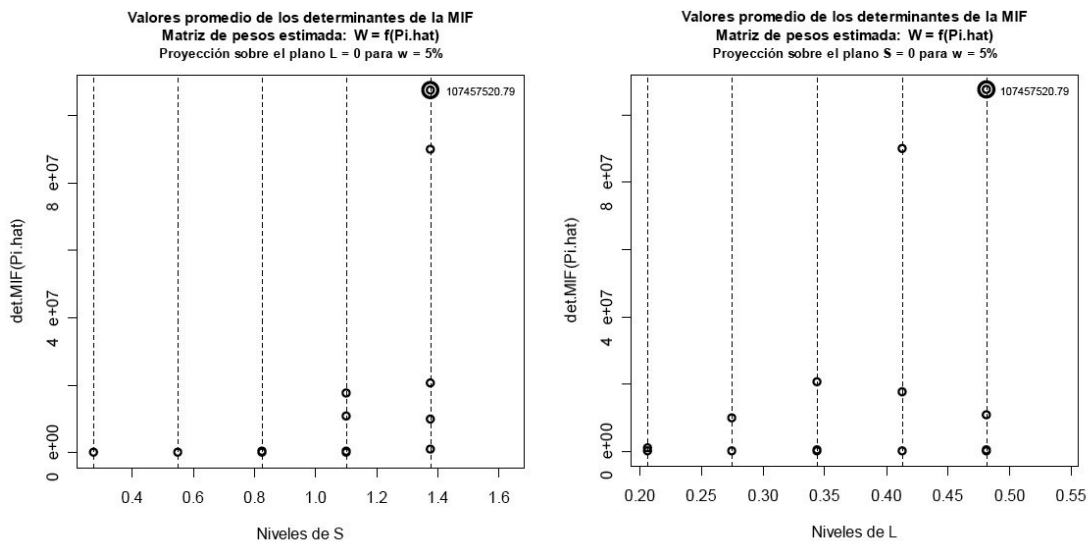


Figura 8.5: Rejilla de los determinantes para todos los casos de  $L$  y  $S$ , estimando  $\pi$  con `pi.hat`.

En las figuras 8.5 y siguientes podemos ver las representaciones gráficas correspondientes para este estadístico, observándose un patrón de comportamiento similar al caso en que calculamos  $\hat{\pi}$  mediante la `pr.obs`: valores altos de  $L$  y de  $S$  parecen dar los mejores resultados para  $w = 5\%$ .



Al determinante de la  $MIF$  calculado en función de  $\mathbf{pi.hat}$ , le aplicamos el logaritmo de la misma manera que hicimos con la primera variante de cálculo y obtenemos la tabla siguiente de datos.

```
> MM.04
```

	S1	S2	S3	S4	S5
L1	0.777830	5.776347	5.890804	4.097636	12.02330
L2	3.103169	7.813191	7.451483	5.282390	14.02557
L3	5.175391	7.277045	6.848491	6.043110	13.86627
L4	6.715430	7.938097	7.263760	16.470200	15.74424
L5	8.051064	9.002338	8.908994	13.708413	16.90321

Podemos ver en las figuras 8.6 y siguientes sus respectivas representaciones gráficas. Como es razonable esperar, la transformación logarítmica tiene la cualidad de estabilizar la variabilidad del estadístico a medida que aumentan los valores de los niveles de las variables  $L$  y  $S$ , lo cual se asemeja razonablemente a los cálculos que realizamos utilizando la `pr.obs` como estimación de  $\pi$ :

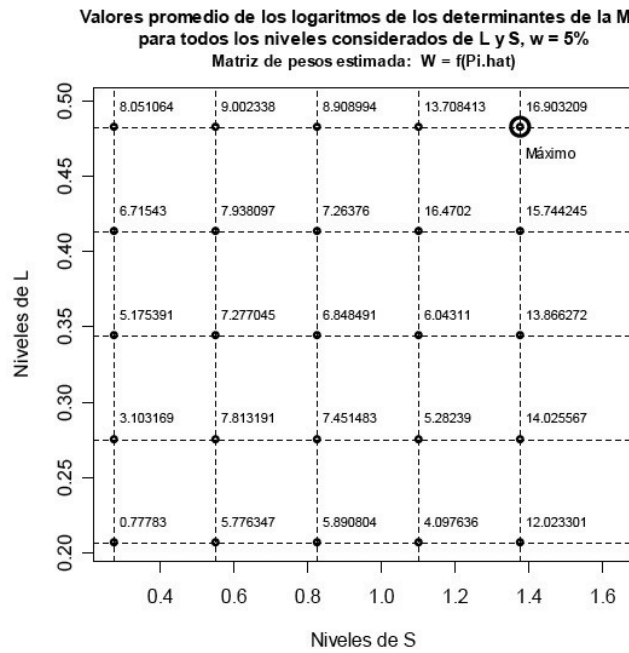
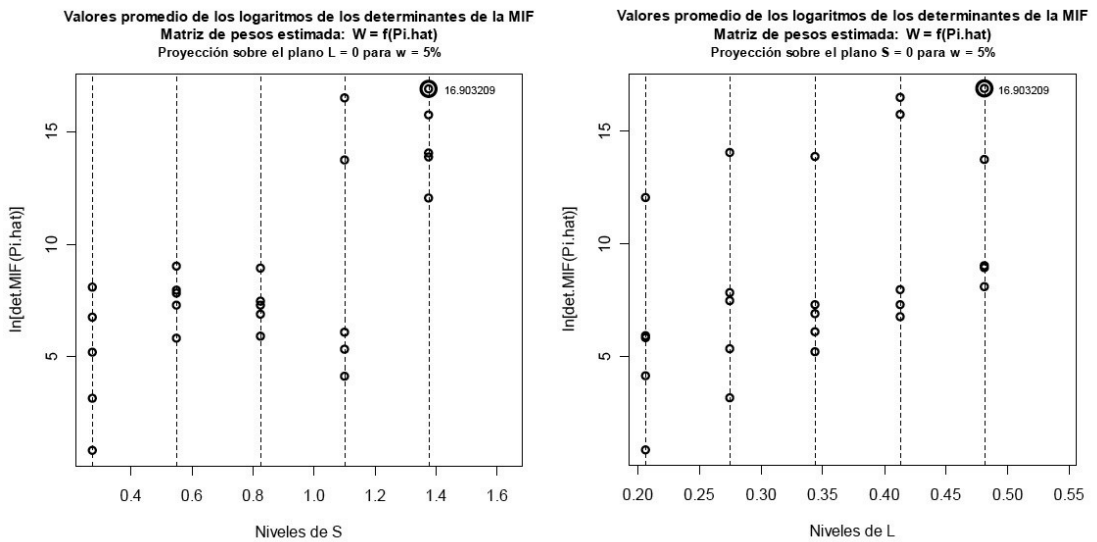


Figura 8.6: Rejilla de los logaritmos de los determinantes para todos los casos de  $L$  y  $S$ , estimando  $\pi$  con  $\hat{\pi}$ .



Del mismo modo que con los determinantes de la  $MIF$  calculados considerando  $pr. obs$ , vemos que los mayores valores se vuelven a observar para los niveles más altos tanto de  $L$  como de  $S$ .

## 8.4. Criterio II: proximidad al máximo

Para ser sistemáticos en la explicación de los resultados obtenidos, repetiremos para el estadístico  $\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$  las mismas gráficas que presentamos para el otro estadístico: gráfico de rejilla y proyecciones sobre ambos planos del estadístico,  $T(S)$  y  $T(L)$ .

Además de evaluar el estadístico en sí mismo, también presentaremos 3 transformaciones del mismo:

- a. la *transformación logit*<sup>6</sup>,
- b. la *transformación angular* o del *arco seno*<sup>7</sup>, y
- c. la *transformación de Freeman-Tukey*<sup>8</sup>.

Desde estos 4 puntos de vista, de amplia aceptación en la literatura sobre el tema —especialmente, los de las 3 transformaciones— podremos ver mejor la repercusión de los resultados al utilizar este criterio propuesto.

### 8.4.1. Probabilidad $\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$

La definición de este estadístico parte de determinar los valores de los factores que hacen máximo el modelo ajustado, los cuales se evalúan en el modelo teórico. Los valores obtenidos se muestran en la matriz siguiente:

```
> MM.05
      S1      S2      S3      S4      S5
L1 0.594798 0.599941 0.794347 0.904871 0.769537
L2 0.734679 0.830373 0.795194 0.876373 0.923440
L3 0.706430 0.786441 0.795721 0.805674 0.927510
L4 0.740100 0.784007 0.821596 0.846119 0.923490
L5 0.820427 0.810562 0.836481 0.931704 0.864095
```

Representamos gráficamente estos valores en las figuras 8.7 y siguientes:

<sup>6</sup>La denominación de esta transformación proviene del acrónimo en inglés “**l**ogarithmic inverse transformation”, que es utilizado por primera vez en: BERKSON, JOSEPH (1944). “Application of the Logistic Function to Bio-Assay”. *Journal of the American Statistical Association*, N°39, pp. 193 *et seq.*, y también en: BERKSON, JOSEPH (1951). “Why I prefer logits to probits”. *Biometrics*, N°7, pp. 327 *et seq.*[Ambos citados en MCCULLAGH Y NELDER (1989)].

<sup>7</sup>Vid. FREEMAN Y TUKEY (1950).

<sup>8</sup>Vid. FREEMAN Y TUKEY (1950) y MILLER, J. J. (1978). “The Inverse of the Freeman-Tukey Double Arcsine Transformation”. *The American Statistician*. N° 32, pp. 138 *et seq.*



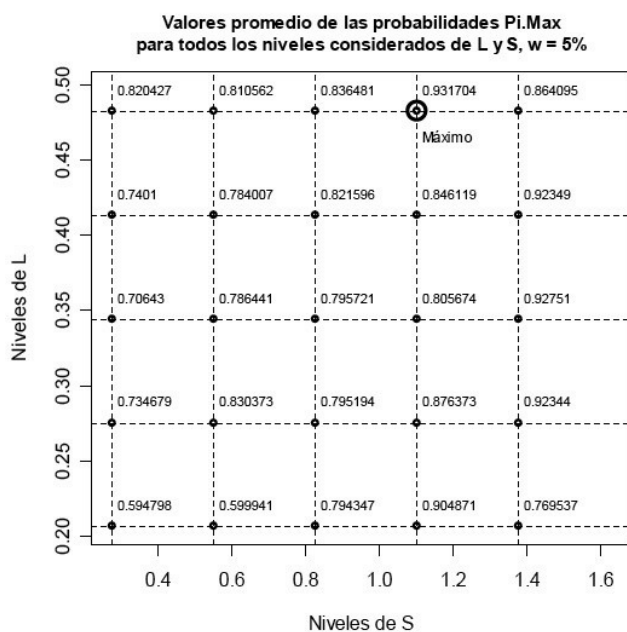
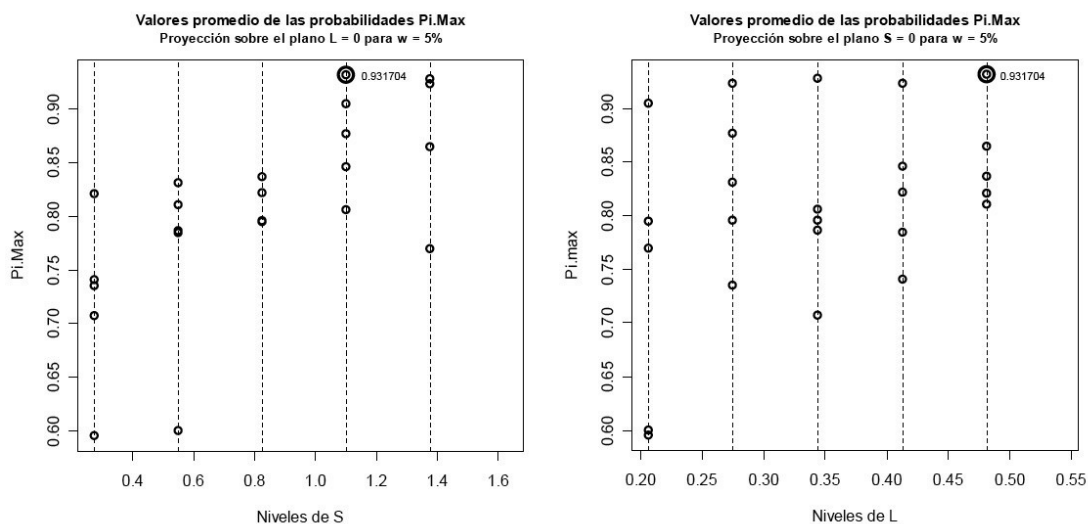


Figura 8.7: Rejilla de probabilidades calculadas en el máximo del modelo ajustado para todos los casos de  $L$  y  $S$ .



Podemos ver que las probabilidades máximas parecen seguir un comportamiento similar que en el caso de los determinantes, observándose los valores mayores para los niveles más altos de las dos variables, aunque con dispersión muy poco estable.

### 8.4.2. Transformación logit de $\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$

Como bien se sabe, la transformación logarítmica inversa o “logit” se define como:

$$\text{logit}[\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})] = \ln \left[ \frac{\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})}{1 - \pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})} \right],$$

la cual evaluaremos en los distintos niveles de  $L$  y  $S$ . Sus valores son:

> MM.06

	S1	S2	S3	S4	S5
L1	0.3838361	0.4052193	1.351330	2.252559	1.205699
L2	1.0184933	1.5882731	1.356523	1.958523	2.490031
L3	0.8781080	1.3036046	1.359762	1.422142	2.549055
L4	1.0464884	1.2891719	1.527198	1.704480	2.490739
L5	1.5192431	1.4536660	1.632275	2.613164	1.849727

Los gráficos correspondientes los podemos ver en las figuras 8.8 y siguientes:

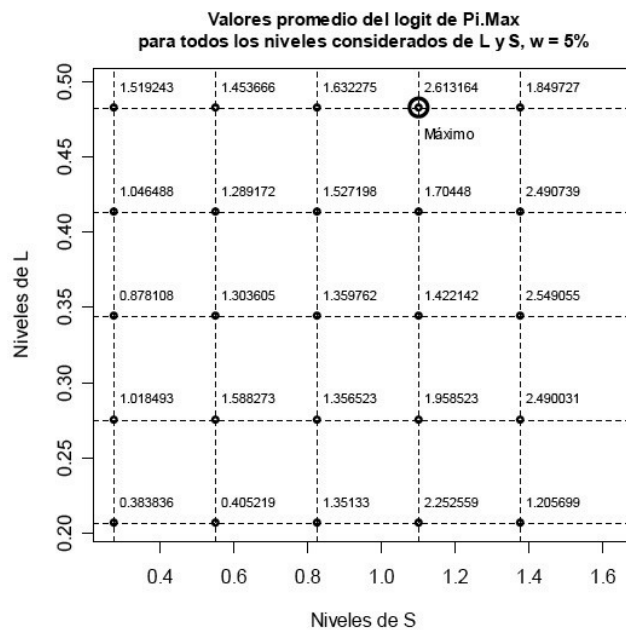
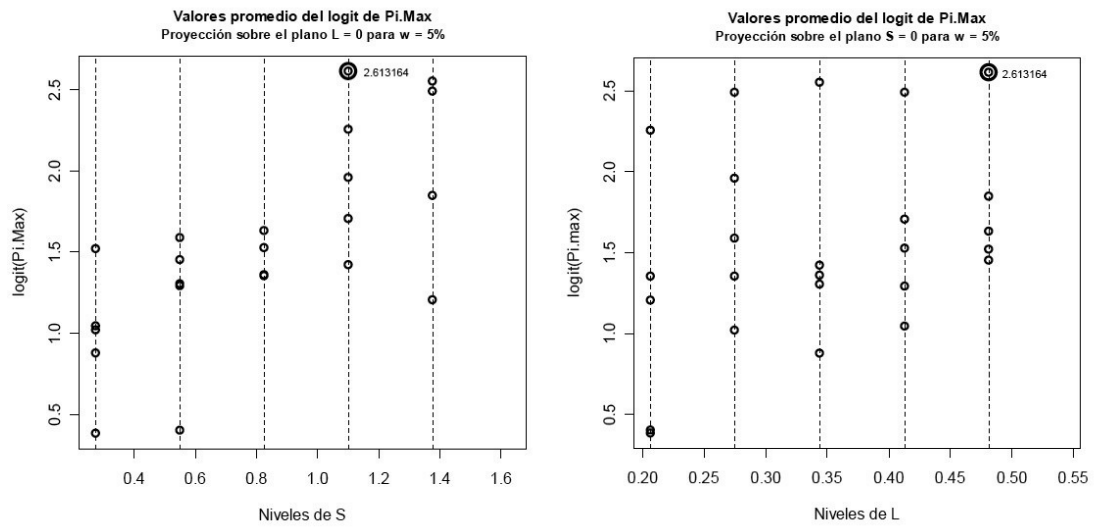


Figura 8.8: Rejilla de probabilidades calculadas en el máximo del modelo ajustado, en escala logit, para todos los casos de  $L$  y  $S$ .



Los resultados de las probabilidades máximas también siguen la misma tendencia que en los casos anteriores, es decir, que se observa para los valores más altos de  $L$  y de  $S$ . Esta transformación también parece estabilizar la variabilidad del estadístico  $\text{logit } \pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$ .

#### 8.4.3. Transformación angular de $\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$

Presentamos a continuación los valores de la transformación del arco seno o transformación angular para continuar nuestra evaluación, que está definida mediante la expresión:

$$\text{asin}[\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})] = \text{arc sen} \left[ \sqrt{\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})} \right]$$

Los resultados obtenidos para cada uno de los valores de la transformación, los mostramos en la tabla que sigue, del mismo modo que como venimos haciendo con las anteriores transformaciones:

> MM.07

	S1	S2	S3	S4	S5
L1	0.8807735	0.8860169	1.100119	1.257254	1.070067
L2	1.0296800	1.1463043	1.101168	1.211510	1.290443
L3	0.9981941	1.0904070	1.101821	1.114280	1.298193
L4	1.0358395	1.0874435	1.134728	1.167691	1.290537
L5	1.1332033	1.1204862	1.154501	1.306391	1.193237

Las representaciones gráficas de esta matriz las vemos en las figuras 8.9 y siguientes:

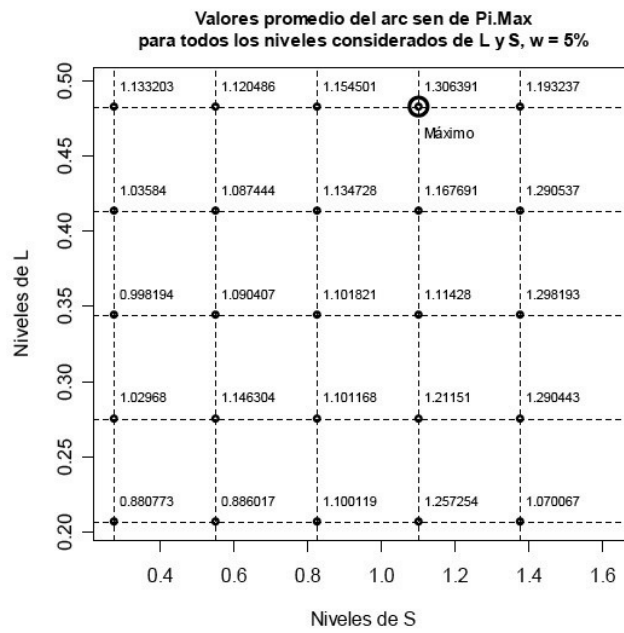
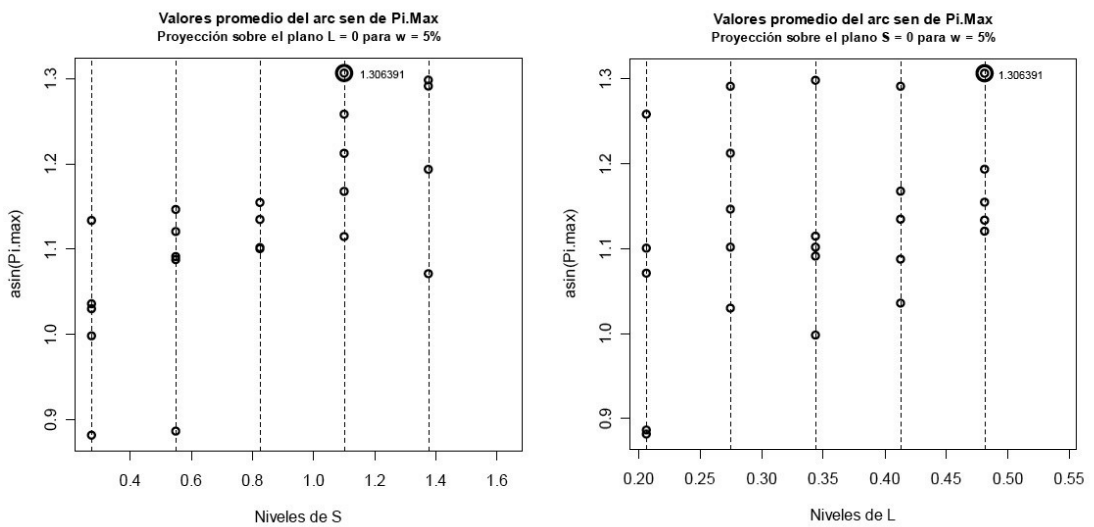


Figura 8.9: Rejilla de probabilidades calculadas en el máximo del modelo ajustado, en escala de la transformación angular, para todos los casos de  $L$  y  $S$ .



El patrón que sigue esta transformación parece ser similar a todas transformaciones que venimos evaluando hasta el momento con ambos criterios.

#### 8.4.4. Transformación de Freeman-Tukey de $\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})$

Finalmente, evaluaremos la transformación Freeman-Tukey:

$$\text{fyt}[\pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})] = \frac{1}{2} \left\{ \text{arc sen} \left[ \sqrt{\frac{m \cdot \pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})}{m+1}} \right] + \text{arc sen} \left[ \sqrt{\frac{[m \cdot \pi(\hat{\mathbf{x}}_{\text{máx}}, \boldsymbol{\beta})] + 1}{m+1}} \right] \right\}$$

Los resultados de evaluar las probabilidades máximas para el valor  $m = 100$  se pueden ver en la siguiente tabla:

> MM.08

	S1	S2	S3	S4	S5
L1	0.8798226	0.8850124	1.096550	1.250578	1.066926
L2	1.0270686	1.1420000	1.097583	1.205947	1.282783
L3	0.9959655	1.0869801	1.098227	1.110496	1.290275
L4	1.0331505	1.0840593	1.130618	1.163007	1.282874
L5	1.1291183	1.1166056	1.150055	1.298188	1.188059

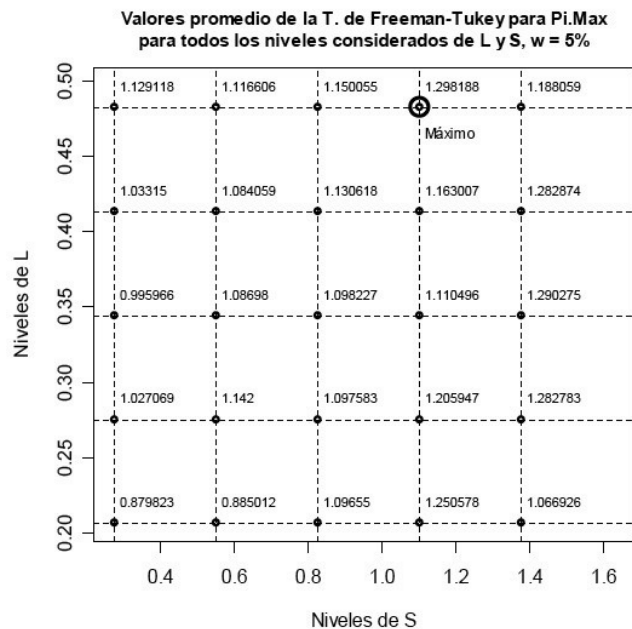
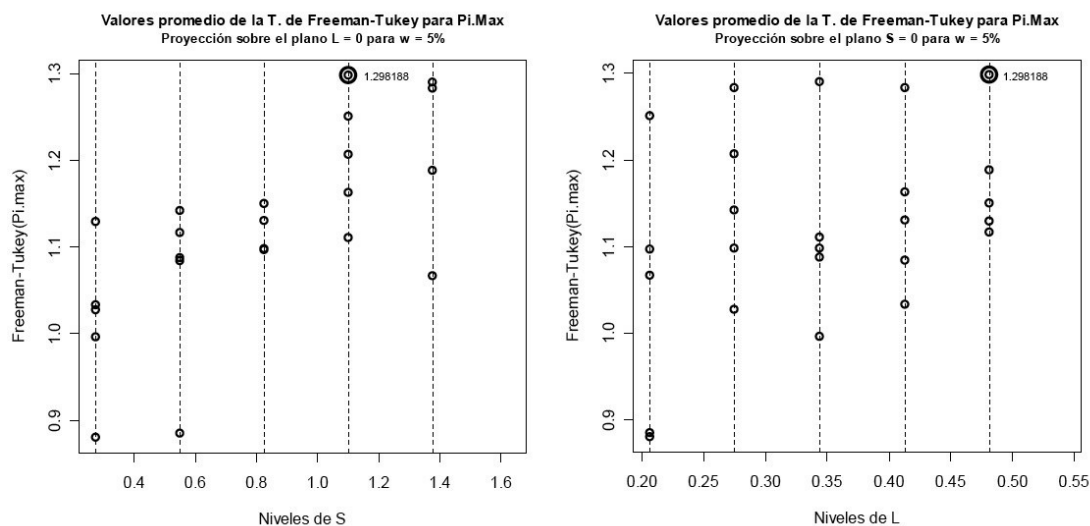


Figura 8.10: Rejilla de probabilidades calculadas en el máximo del modelo ajustado, en escala de la transformación Freeman-Tukey, para todos los casos de  $L$  y  $S$ .

En las figuras 8.10 y siguientes, las representaciones gráficas de estos resultados:



Los resultados que podemos ver para la transformación Freeman-Tukey de la probabilidad, siguen la misma tendencia que todos los casos de transformaciones que hemos evaluado.

## 8.5. Algunas observaciones

Un cuadro que nos resulta relativamente útil para resumir la información de todos los estadísticos de medida de calidad de ajuste para las 25 filas de la matriz  $\mathbf{LS}$  es el de la figura 8.11, en la página siguiente.

Podemos ver que, para el caso de los determinantes calculados en sus dos modalidades<sup>9</sup>, parece bastante claro que la maximización de los mismos se consigue cuando ambas variables,  $L$  y  $S$ , toman los niveles más altos de aquellos que hemos considerado. El caso del logaritmo del determinante de la  $MIF$ , cuando la matriz de “pesos” la calculamos en función de las probabilidades observadas, presenta su máximo para el cuarto nivel de la variable  $L$ , aunque la diferencia que existe con el quinto valor (ver fig. 8.4) no parece ser demasiado alta. Sin mucha exageración, podríamos conjeturar que los niveles más altos de  $L$  y  $S$  conducirán a los máximos valores de los promedios de los determinantes de las  $MIF$  de los 15 cuadros de puntos, ya sean en escala natural o bien en su transformación a escala logarítmica. Acerca de esta última, puede verse en todas las proyecciones sobre los planos  $L = 0$  y  $S = 0$  que se ha reducido la variabilidad de los puntos observados con respecto a la que tienen los valores en escala natural, lo cual era de natural de esperar.

<sup>9</sup>Es decir, calculados con las dos probabilidades estimadas que venimos comentando desde secciones anteriores, tal como aparecen en las ecuaciones (7.1) y (7.2).

Cuadro comparativo de los resultados obtenidos para cada estadístico utilizado,  $w = 5\%$ 

CONDICIONES EXPERIMENTALES		FRECUENCIAS									
PARÁMETRO	NIVELES	delMIF $W = f(\text{pr. obs})$	ln(delMIF) $W = f(\text{pr. obs})$	delMIF $W = f(\text{Pl. obs})$	ln(delMIF) $W = f(\text{Pl. obs})$	Pl. Max	Tr. logit Pl. Max	Tr. angular Pl. Max	Tr. F-T Pl. Max		
L (todo del factorial)	L1 0.2064356										
	L2 0.2752474										
	L3 0.3440593										
	L4 0.4128711	✓									
	L5 0.4816830	✓		✓		✓	✓	✓	✓		
S (salto entre centros)	S1 0.2752474										
	S2 0.5504948										
	S3 0.8257422										
	S4 1.1009896					✓	✓	✓	✓		
	S5 1.3762370	✓	✓	✓	✓	0.931704	2.613164	1.306391	1.298188		
VALOR MÁXIMO		164539317	18.31340	107457520.8	16.90321	0.931704	2.613164	1.306391	1.298188		

Figura 8.11: Cuadro de resumen para todos los estadísticos de medida de calidad de ajuste evaluados en todos los niveles de  $L$  y  $S$ .

El caso del criterio II de la proximidad al máximo —calculadas en el máximo del modelo ajustado y sus transformaciones— no parece evidenciarse una diferencia muy sustancial con respecto del caso de los determinantes de la  $MIF$  —criterio I— salvo por el detalle que los valores máximos de los estadísticos de valoración no se consiguen para el máximo nivel de  $S$  sino para el inmediato anterior. Sin consideraciones de tamaños de muestra mayores para cada punto de diseño ( $m \gg 100$ ), no tenemos evidencia que la diferencia entre dos valores “relativamente grandes” y próximos entre sí, resulten estadísticamente significativos. No obstante ello, al observar las proyecciones de los mismos sobre los planos  $L = 0$  y  $S = 0$ , también podemos notar que no hay diferencias muy marcadas con respecto a los valores calculados en el nivel más alto de  $S$ .

Finalmente, y ordenando ideas, creemos razonable indicar que los mejores valores de las probabilidades (criterio I) parece que se obtienen en condiciones muy similares a las de los determinantes de las  $MIF$  (criterio II), en el sentido en que valores altos de ambas variables parecen hacer que ambos estadísticos —y sus transformaciones— tendieran hacia los máximos valores respectivos dentro de los niveles considerados de ambas variables de estudio.

No obstante todas estas observaciones que hemos realizado, las mismas se refieren al problema “incompleto” en el sentido en que hemos introducido niveles en dos variables y la tercera la hemos dejado constante. Para completar el problema, queda por agregar que la inclusión de niveles para la tercera de las variables de estudio,  $w$ , permitirá visualizar de forma total todas las variables de estudio — $w$ ,  $L$  y  $S$ —y todos sus niveles —5 por variable— de forma que tendremos un total de  $5 \times 5 \times 5 = 125$  casos de estudio, entre los que averiguaremos cuáles de estos niveles de las variables son los que maximizan los dos criterios que hemos propuesto para evaluar los ajustes. De esta manera, éste sería el último paso en nuestro estudio —previo a las conclusiones— con el que intentaremos cumplir con el objetivo que nos propusimos al iniciar este trabajo.