UNAP
FINESI

IV Semestre

Modelos Discretos

Regresión Multinomial

Mtr. Alcides Ramos Calcina

Regresión Multinomial

Validación del Modelo

Mtr. Alcides Ramos Calcina

Validación del modelo

Realizaremos la validación para los 5 modelos estimados.

En cuanto a la validación del modelo, el test de Wall es equivalente al intervalo de confianza que podemos calcular para el parámetro β_{is}.

Si se calcula un intervalo de confianza al 95% para los parámetros, ¿qué tendríamos que mirar para saber si son estadísticamente significativo? Que el cero esté en el intervalo.

$$IC: P\left[\hat{\beta}_{1s} - z_{1-\alpha/2}\widehat{SE}_s\left(\hat{\beta}_{1s}\right) \le \beta_{1s} \le \hat{\beta}_{1s} + z_{1-\alpha/2}\widehat{SE}_s\left(\hat{\beta}_{1s}\right)\right] = 1-\alpha$$

Ejemplo 5

- **Modelo 1**

Continuando con el ejemplo 5, el intervalo de confianza para las categorías es:

Para β_{11}:

$$IC: P\left[1.25653-(1.96)(0.37467)\leq \beta_{11}\leq 1.25653+(1.96)(0.37467)\right]=0.95$$

$$IC: P\left[0.52220\leq \beta_{11}\leq 1.99086\right]=0.95$$

Para β_{12}:

$$IC: P\left[1.00938-(1.96)(0.42751)\leq \beta_{12}\leq 1.00938+(1.96)(0.42751)\right]=0.95$$

$$IC: P\left[0.17148\leq \beta_{12}\leq 1.84729\right]=0.95$$

Ejemplo 5

```
confint(modelo1, level = 0.95)
, , 1

                  2.5 %     97.5 %
(Intercept) -1.2012893 -0.7006696
hist1        0.5222045  1.9908576

, , 2

                  2.5 %     97.5 %
(Intercept) -1.5305447 -0.9704159
hist1        0.1714807  1.8472880
```

Se observa que los β_{is} no pude tomar el valor de cero, en ninguna ecuación, por tanto, podemos decir que la variable ***hist*** es estadísticamente significativa.

Ejemplo 5

Test de la razón de verosimilitud

Al momento de estimar los coeficientes, no se tiene los *p*-valor para los coeficientes, si queremos saber si la variable explicativa es significativa, utilizamos el test de la razón de verosimilitud, en esta comparamos el modelo nulo con el modelo que si tiene la variable:

```
modelo0 <- multinom(me ~ 1)
# weights:  6 (2 variable)
initial  value 452.628263
final  value 402.599009
converged
```

El valor de probabilidad asociado al estadístico Chi-cuadra $Pr(chi) = 0.001614001 < 0.05$, entonces rechazamos la hipótesis nula de que $\beta_1 = 0$. Es decir, que la variable **hist** es estadísticamente significativa.

```
anova(modelo1, modelo0)
  Model Resid. df Resid. Dev   Test    Df LR stat.     Pr(Chi)
1     1       822   805.1980           NA       NA          NA
2  hist       820   792.3399 1 vs 2     2 12.85808 0.001614001
```

Ejemplo 5

- **Modelo 2**

Como se analizo anteriormente, nos interesa interpretar aquellos betas que son distintos de cero. Y puede pasar que alguno de ellos sea significativo pero otros no, por consiguiente solicitemos la estimemos el intervalo de confianza para los parámetros.

```
IC_dect <- confint(modelo2, level = 0.95)
IC_dect[,,1] # IC: Ultimo año / Nunca
                  2.5 %      97.5 %
(Intercept) -4.5993833 -0.5310183
dect1       -1.4169276  2.8294454
dect2        0.0551994  4.1572913
```

Esto quiere decir que cuando se compara el ultimo año vs nunca, de $dect = 0$ a $dect = 1$ no hay evidencias estadísticamente significativas.

Se obtuvo los intervalos de confianza para los betas de la primera ecuación. El intervalo de confianza para **dect1** el beta podría ser cero.

$$IC : P\left[-1.41693 \leq \beta_{11\_dect1} \leq 2.82945\right] = 0.95 \qquad 0 \in \left[-1.41693;\ 2.82945\right]$$

Ejemplo 5

Sin embargo, si que las hay cuando comparo $dect = 0$ con $dect = 2$, como se muestra

$$IC : P\left[0.05520 \leq \beta_{11\_dect2} \leq 4.15729\right] = 0.95$$

Por tanto, $\quad 0 \notin \left[0.05520;\ 4.15729\right]$

¿Cómo lo interpreto? el odds de hacerse una mamografía dentro del ultimo año entre las mujeres que piensan que es muy probable que la mamografía detecte un nuevo caso de cáncer ($dect = 2$), es 8.22 veces mayor que el odds entre las mujeres que piensan que no es probable que la mamo lo detecte ($dect = 1$).

Ejemplo 5

Tendríamos que hacer lo mismo si queremos comparar el logit de más de un año vs nunca.

```
IC_dect[,,2] # IC: Más de un año / Nunca
                  2.5 %      97.5 %
(Intercept) -2.2992503 -0.0579843
dect1       -1.6359055  0.8506970
dect2       -0.9656586  1.3612557
```

En este caso, los intervalos de confianza de los betas, ambos incluyen al cero.

Por lo tanto, ninguno de esos 2 betas es estadísticamente significativo.

Lo cual quiere decir que de los 4 betas que hemos estimado solo tenemos 1 que es estadísticamente significativo.

Ejemplo 5

Por consiguiente, tenderemos que evaluar la variable a través de la **razón de verosimilitud**.

```
modelo0 <- multinom(me ~ 1)
# weights:  6 (2 variable)
initial  value 452.628263
final  value 402.599009
converged
anova(modelo2, modelo0)
  Model Resid. df Resid. Dev   Test    Df LR stat.      Pr(Chi)
1     1       822   805.1980           NA       NA           NA
2  dect       818   778.4011 1 vs 2     4 26.79694 2.184912e-05
```

El valor de probabilidad asociado al estadístico Chi-cuadra $Pr(chi) = 0.00002 < 0.05$, entonces rechazamos la hipótesis nula de que $\beta_2 = 0$. Es decir, que la variable **dect** en conjunto es estadísticamente significativa.

Ejemplo 5

A modo de resumen mostramos el AIC y la razón de verosimilud de los cinco modelos individuales, para determinar su significancia.

Razón de verosimilitud y significancia

Modelo	AIC	Razón de Verosimilitud Pr.(Chi)	Sig. (< 0.05)
modelo1	800.34	0.00161	**
modelo2	709.40	0.00002	***
modelo3	757.41	0.00000	***
modelo4	777.94	0.00000	***
modelo5	796.06	0.00019	***

De acuerdo a la tabla, todos los modelos univariantes son significativos, por tanto, ahora incluiremos todas las variables en una sola y generamos un modelo multivariante.

Regresión Multinomial

Modelo Múltiple

Mtr. Alcides Ramos Calcina

Modelo de Regresión Multinomial Múltiple

- **Modelo 6**

$$\log\left[\frac{P(Y=s)}{P(Y=0)}\right] = \beta_{0s} + \beta_{1s}hist + \beta_{2s}dect + \beta_{3s}symp + \beta_{4s}bse + \beta_{5s}pb$$

Continuando con el ejemplo 5, mostramos la estimación del modelo 6.

```
modelo6 <- multinom(me ~ hist + dect + symp + bse + pb, data = datos)
# weights:  30 (18 variable)
initial  value 452.628263
iter  10 value 348.174100
iter  20 value 346.951267
final  value 346.950964
converged
```

Ejemplo 5

```
summary(modelo6)
Call:
multinom(formula = me ~ hist + dect + symp + bse + pb, data = datos)

Coefficients:
  (Intercept)    hist1       dect1      dect2      symp2     symp3
1  -2.9990607 1.366274  0.01700706  0.9041717  0.1101408 1.9249158
2  -0.9860258 1.065446 -0.92448094 -0.6906153 -0.2901346 0.8172916
     symp4     bse1         pb
1 2.457230 1.291772 -0.2194421
2 1.132224 1.052183 -0.1482079

Std. Errors:
  (Intercept)     hist1     dect1     dect2     symp2     symp3
1    1.539293 0.4375239 1.1619394 1.1268643 0.9228324 0.7776651
2    1.111829 0.4593986 0.7137332 0.6871025 0.6440622 0.5397872
     symp4      bse1         pb
1 0.775400 0.5299080 0.07551477
2 0.547666 0.5149934 0.07636886

Residual Deviance: 693.9019
AIC: 729.9019
```

Ejemplo 5

Para determinar la significancia de las variables en el modelo utilizamos el intervalo de confianza para cada beta.

```
confint(modelo6, level = 0.95)
, , 1
                 2.5 %      97.5 %
(Intercept) -6.0160192  0.01789780
hist1        0.5087428  2.22380478
dect1       -2.2603524  2.29436653
dect2       -1.3044418  3.11278514
symp2       -1.6985774  1.91885907
symp3        0.4007202  3.44911153
symp4        0.9374738  3.97698584
bse1         0.2531716  2.33037280
pb          -0.3674483 -0.07143587
```

Ejemplo 5

```
, , 2

                  2.5 %      97.5 %
(Intercept) -3.16516983 1.193118231
hist1        0.16504089 1.965850477
dect1       -2.32337238 0.474410491
dect2       -2.03731145 0.656080761
symp2       -1.55247327 0.972204160
symp3       -0.24067187 1.875255164
symp4        0.05881781 2.205629242
bse1         0.04281458 2.061551754
pb          -0.29788816 0.001472263
```

Ejemplo 5

Analicemos los intervalos de cada variable:

- **hist** : en las dos ecuaciones $\beta 1s \neq 0$, por tanto, la variable es significativa
- **dect** : Vemos que ninguno de los cuatro coeficientes parece ser significativo, por lo que esta variable es no significativa, por lo que, se ajustara el modelo sin esta variable.
- **symp** : en la ecuación el primer beta es no significativo y en la segunda ecuación los betas 1 y 2 son no significativos, por lo que, podríamos simplificar la codificación de la variable agrupando los dos primeros niveles. En general, la variable es significativa.
- **bse** : en las dos ecuaciones sus betas son diferentes de cero, por tanto, la variable es significativa.
- **pb** : en la ecuación 1 es significativa y la ecuación 2 no lo es, por tanto, consideramos como significativa a la variable.

Ejemplo 5

- **Modelo 7**

A continuación, ajustaremos un nuevo modelo, en el cual excluiremos la variable **dect**.

$$\log\left[\frac{P(Y=s)}{P(Y=0)}\right]=\beta_{0s}+\beta_{1s}hist+\beta_{2s}symp+\beta_{3s}bse+\beta_{4s}pb$$

```
modelo7 <- multinom(me ~ hist + symp + bse + pb, data = datos)
# weights:  24 (14 variable)
initial  value 452.628263
iter  10 value 351.526518
final  value 351.173900
converged
```

Ejemplo 5

```
summary(modelo7)
Call:
multinom(formula = me ~ hist + symp + bse + pb, data = datos)

Coefficients:
  (Intercept)    hist1      symp2     symp3    symp4      bse1
1   -2.181670 1.373809  0.1530116 2.0855801 2.618928 1.2666132
2   -1.733803 1.115519 -0.3074208 0.8418555 1.142751 0.9816765
          pb
1 -0.2502638
2 -0.1380179

Std. Errors:
  (Intercept)     hist1     symp2     symp3     symp4      bse1
1   1.0256060 0.4339476 0.9162843 0.7706965 0.7679758 0.5245529
2   0.8727627 0.4584496 0.6445228 0.5362119 0.5415082 0.5084326
          pb
1 0.07407661
2 0.07362492

Residual Deviance: 702.3478
AIC: 730.3478
```

Ejemplo 5

Para verificar que la exclusión de la variable **dect** calcularemos la razón de verisimilitud entre los modelos 6 y 7, planteando la siguiente hipótesis:

$$H_0\text{: Vector de betas} = 0$$
$$H_1\text{: Vector de betas} \neq 0$$

De los resultados, se tiene que el p-valor = 0.07655 > 0.05, entonces aceptamos H_0, quiere decir que el vector de betas es igual a cero; por tanto, no entra la variable **dect** en el modelo al no ser significativa.

```
anova(modelo6, modelo7)
                         Model Resid. df Resid. Dev    Test     Df
1         hist + symp + bse + pb       810   702.3478             NA
2 hist + dect + symp + bse + pb       806   693.9019 1 vs 2      4
  LR stat.     Pr(Chi)
1        NA          NA
2 8.445871 0.07654504
```

Uso de la función *step*

Partiendo del modelo grande que nosotros hemos introducido, para ver cual seria mejor y el criterio que utiliza para ver cual es el mejor, es el valor del AIC.

```
step(modelo6)
Start:  AIC=729.9
me ~ hist + dect + symp + bse + pb

trying - hist
# weights:  27 (16 variable)
initial  value 452.628263
iter  10 value 353.888326
iter  20 value 352.579296
final  value 352.579290
converged
.
.
.
```

Uso de la función *step*

```
       Df      AIC
<none> 18 729.9019
- dect 14 730.3478
- bse  16 735.5729
- pb   16 736.0389
- hist 16 737.1586
- symp 12 756.3230
Call:
multinom(formula = me ~ hist + dect + symp + bse + pb, data = datos)

Coefficients:
  (Intercept)    hist1       dect1      dect2      symp2     symp3    symp4     bse1         pb
1  -2.9990607 1.366274  0.01700706  0.9041717  0.1101408 1.9249158 2.457230 1.291772 -0.2194421
2  -0.9860258 1.065446 -0.92448094 -0.6906153 -0.2901346 0.8172916 1.132224 1.052183 -0.1482079

Residual Deviance: 693.9019
AIC: 729.9019
```

Si no le quitamos ningún parámetro al modelo, tenemos un AIC de 729.9019, que resulta ser el más pequeño de todos. Así mismo, el que le sigue el modelo sin la variable **dect**, pues su AIC = 730.3478, por tanto, se recomienda quitar esta variable del modelo como ya se verifico anteriormente.

Ejemplo 5

Una ves elegido nuestro modelo múltiple (modelo 7), estimamos los OR:

```
exp(coef(modelo7))
  (Intercept)    hist1     symp2    symp3     symp4     bse1        pb
1   0.1128529 3.950369 1.1653385 8.049259 13.721009 3.548813 0.7785954
2  0.1766115 3.051151 0.7353411 2.320669  3.135382 2.668927 0.8710831
```

- En el caso de la variable síntoma (**symp**), las mujeres que están muy en desacuerdo con la frase: "La mamografía no es necesaria a menos que aparezca algún síntoma" tienen un odds 13.72 veces mayor de haberse hecho una mamografía el último año, y un odds 3.14 veces mayor de habérsela hacho hace más de un año, comparado con mujeres que no están muy deacuerdo con esa frase.

Ejemplo 5

- Las mujeres con historia familiar de cáncer tienen un odds 3.9 veces mayor de haber hecho uso de la mamografía el último año y 3.1 veces mayor al hacerlo hace más de un año, comparado con mujeres que no tienen un familiar con historia de cáncer.

- Haber aprendido a explorarse el pecho es un factor significativo a la hora de haberse hecho una mamografía en el último año, pero no lo es a la hora de hacerlo hace más de un año.

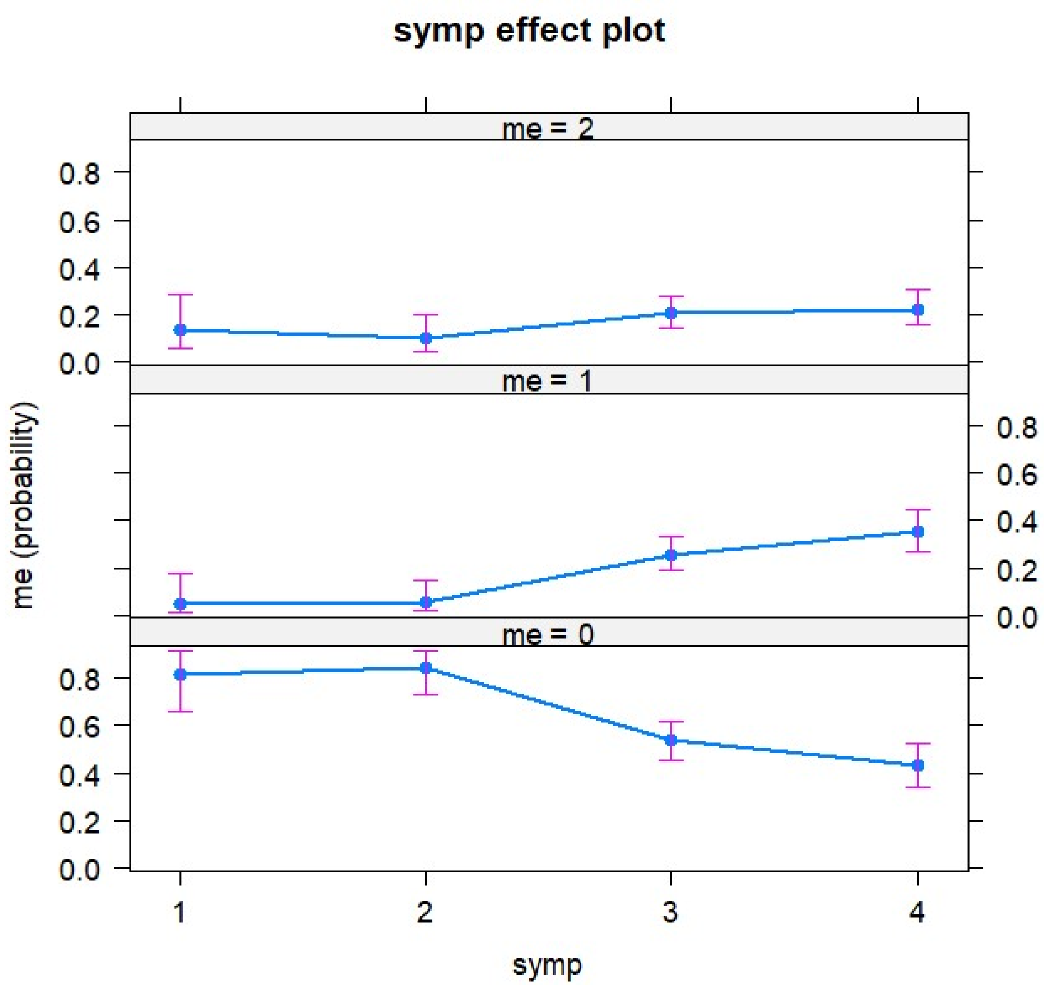
- La variable beneficio percibido (pb) tiene un OR < 1 ya que valores más altos indica creencia de menor beneficio, además, cuanto más aumenta la desconfianza más disminuye el odds de someterse a una mamografía ya sea reciente o no recientemente.

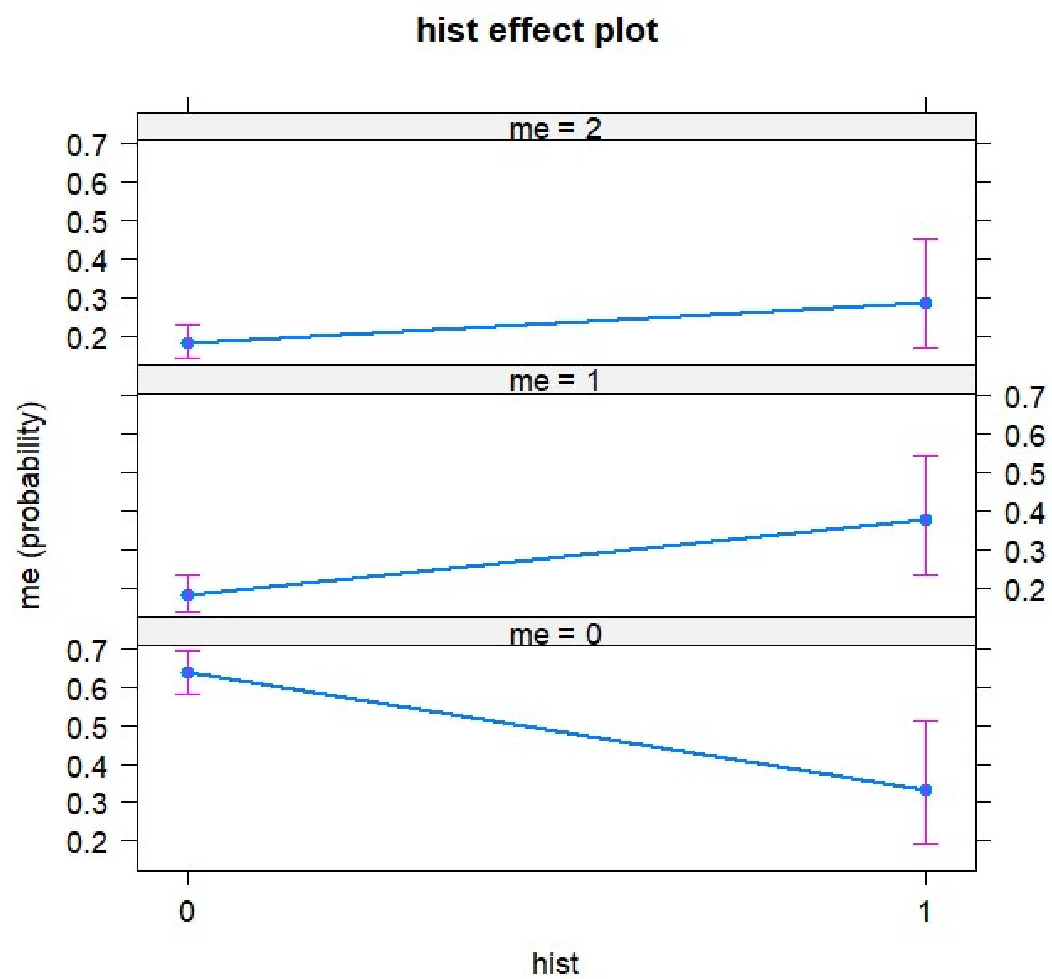
Ejemplo 5

Podemos ver como cambia la probabilidad de frecuencia del chequeo para las distintas variables de forma gráfica:

```
# install.packages("effects")
# library(effects)
plot(effect("symp", modelo7))
plot(effect("hist", modelo7))
plot(effect("bse", modelo7))
plot(effect("pb", modelo7))
```

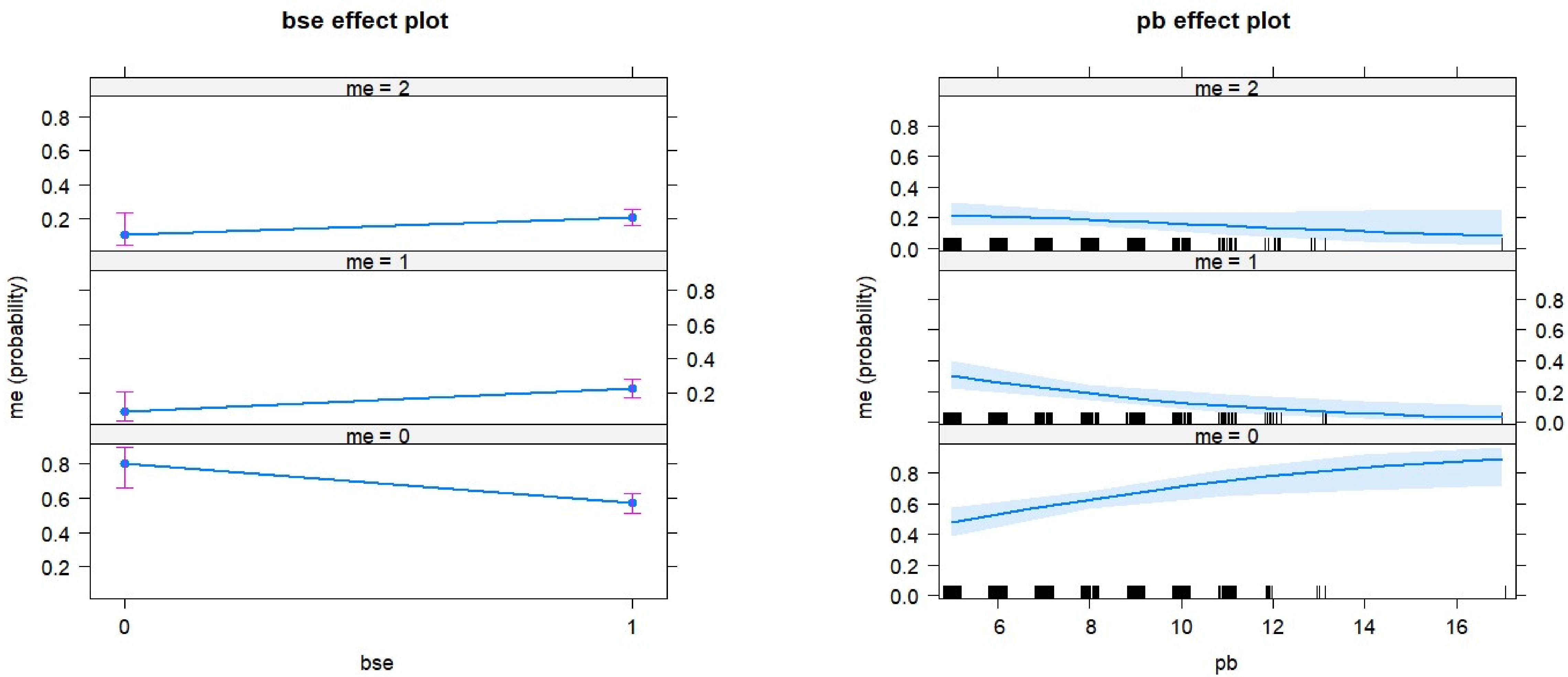
Ejemplo 5





Mtr. Alcides Ramos Calcina

Ejemplo 5



Mtr. Alcides Ramos Calcina

FINESI

Modelos Discretos

IV Semestre

GRACIAS

https://aulavirtual2.unap.edu.pe/