

Examen Parcial 1

Ana B. Coronel, 107874

11 de noviembre de 2019

El sistema de salud en México cuenta con una base de datos gestacionales, es decir, características de mujeres embarazadas y del producto. La base cuenta con 535 casos de embarazo descritos en 12 variables:

1. Anestesia: Indicador si durante el parto se usó anestesia (1-no, 2-sí)
2. Complica: Indicador si la madre tuvo algún tipo de complicación durante su embarazo (1-no, 2-sí)
3. Edad: Edad de la madre en años
4. Escolaridad: Escolaridad de la madre (1-primaria, 2-secundaria, 3-preparatoria, 4-universidad)
5. Estado_civil: Estado civil de la madre (1-casada, 2-divorciada, 3-soltera, 4-uniónlibre)
6. Forma_nac: Tipo de parto (1-cesárea, 2-vaginal)
7. Hospital: Nombre del hospital en donde se atendió el embarazo (1-Ángeles, 2-Español, 3- General, 4-Médica Sur, 5-Oaxaca, 6-Pemex Sur, 7-Tlaxcala)
8. Ind_cesarea: Motivo por el que se llevó acabo la cesárea (1-DCP (desproporción céfalo pélvica), 2-electiva, 3-emergencia fetal, 4-emergencia materna, 5-iterativa, 6-no hubo)
9. Monitoreo: Indicador si la madre recibió un monitoreo fetal durante el embarazo (1-no, 2-sí)
10. Mortalidad: Mortalidad del producto (1-no, 2-sí)
11. Peso_madre: Peso de la madre al inicio del embarazo (en Kgs)
12. Peso_prod: Peso al nacer del producto (en Kgs)

A continuación, se muestran las primeras seis observaciones la base de datos *Gestacional*. Por la descripción proporcionada de las variables, se puede decir que las únicas variables numéricas son edad, peso de la madre y peso del producto; la escolaridad es una variable ordinal y el resto son nominales.

##	Anestesia	Complica	Edad	Escolaridad	Estado_civil	Forma_nac	Hospital
## 1	2	1	37	4	1	1	1
## 2	1	2	22	3	4	2	7
## 3	2	1	31	1	4	1	5
## 4	2	2	24	2	4	1	7
## 5	2	2	39	4	1	1	1
## 6	1	2	19	1	1	2	5

##	Ind_cesarea	Monitoreo	Mortalidad	Peso_madre	Peso_prod
## 1	2	2	1	60	2.75
## 2	6	2	1	50	2.40
## 3	1	2	1	53	3.30
## 4	5	2	1	99	3.71
## 5	5	2	1	70	3.41
## 6	6	2	1	65	3.25

Análisis exploratorio de datos

Como se observa en el siguiente resumen, la base de datos no tiene observaciones faltantes. Adicional, se observa que:

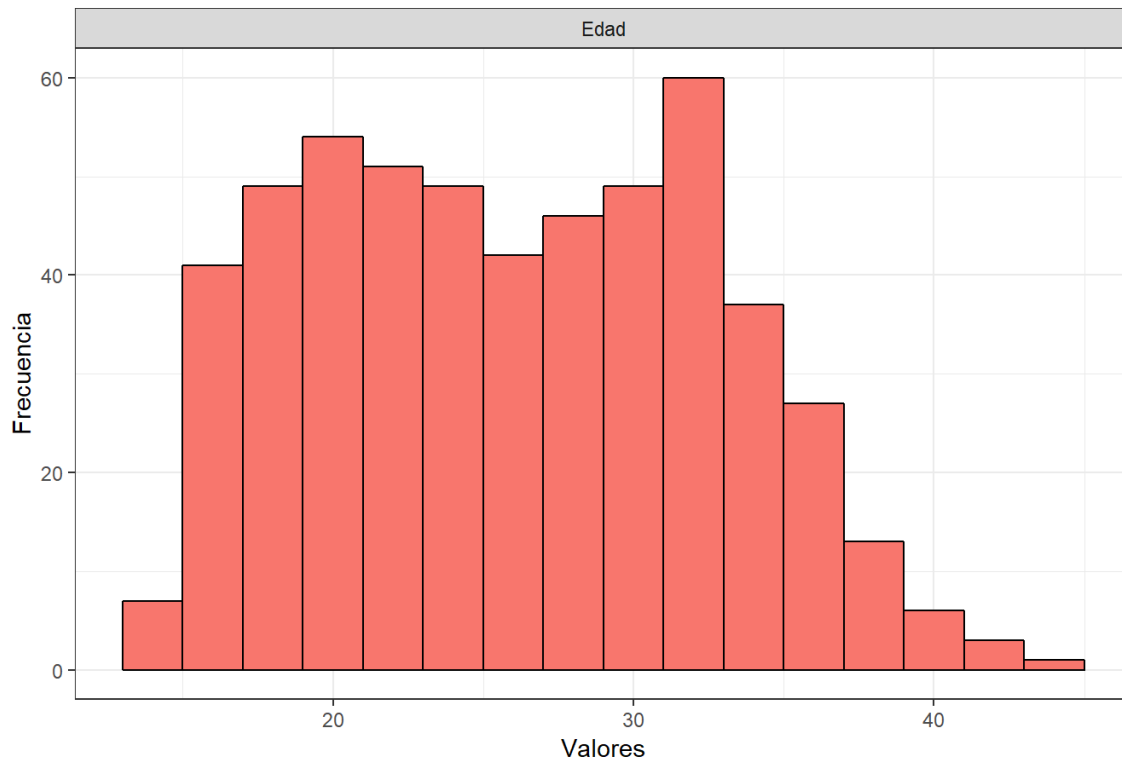
- En la mayoría de los casos de embarazo, se utilizó anestesia durante el parto
- La edad mediana de las madres es de 26 años. La edad mínima es de 14 y la máxima de 44
- Escolaridad: únicamente 158 madres tienen universidad
- La mayoría de los nacimientos ocurrieron con cesárea
- Dos de los hospitales (3- Hospital General y 5-Oaxaca) concentran más de la mitad de los nacimientos 55.8878505 %.

- Únicamente murieron 35 niños
- En más del 75% (464) de los casos, las madres recibieron monitoreo fetal durante el embarazo
- El peso de las madres toma valores entre 37 y 108 kg. El peso mediano es de 58 kg.
- El peso de los productos toma valores entre 0.004 kg y 4.5 kg. El rango intercuartílico es de 0.619 kg.

```
## Anestesia Complica      Edad      Escolaridad Estado_civil Forma_nac
## 1:118      1:253  Min.   :14.00   1:100      1:256      1:309
## 2:417      2:282  1st Qu.:21.00   2:186      2:   6      2:226
##                               Median :26.00   3:  91      3:  80
##                               Mean   :26.53   4:158      4:193
##                               3rd Qu.:32.00
##                               Max.   :44.00
##
## Hospital Ind_cesarea Monitoreo Mortalidad  Peso_madre
## 1: 76      1: 72      1: 71      1:500      Min.   : 37.00
## 2: 60      2: 60      2:464      2: 35      1st Qu.: 52.00
## 3:163      3: 94
## 4: 33      4: 37
## 5:136      5: 46
## 6:  2      6:226
## 7: 65
##
## Peso_prod
## Min.   :0.004
## 1st Qu.:2.658
## Median :2.980
## Mean   :2.876
## 3rd Qu.:3.277
## Max.   :4.500
##
```

En la siguiente gráfica, se muestra la distribución de la edad de las madres. Existen pocas observaciones en edades mayores a 40 años. La distribución tiene dos modas: una en la edad aproximada de 20 años y otra en la edad aproximada de 32 años

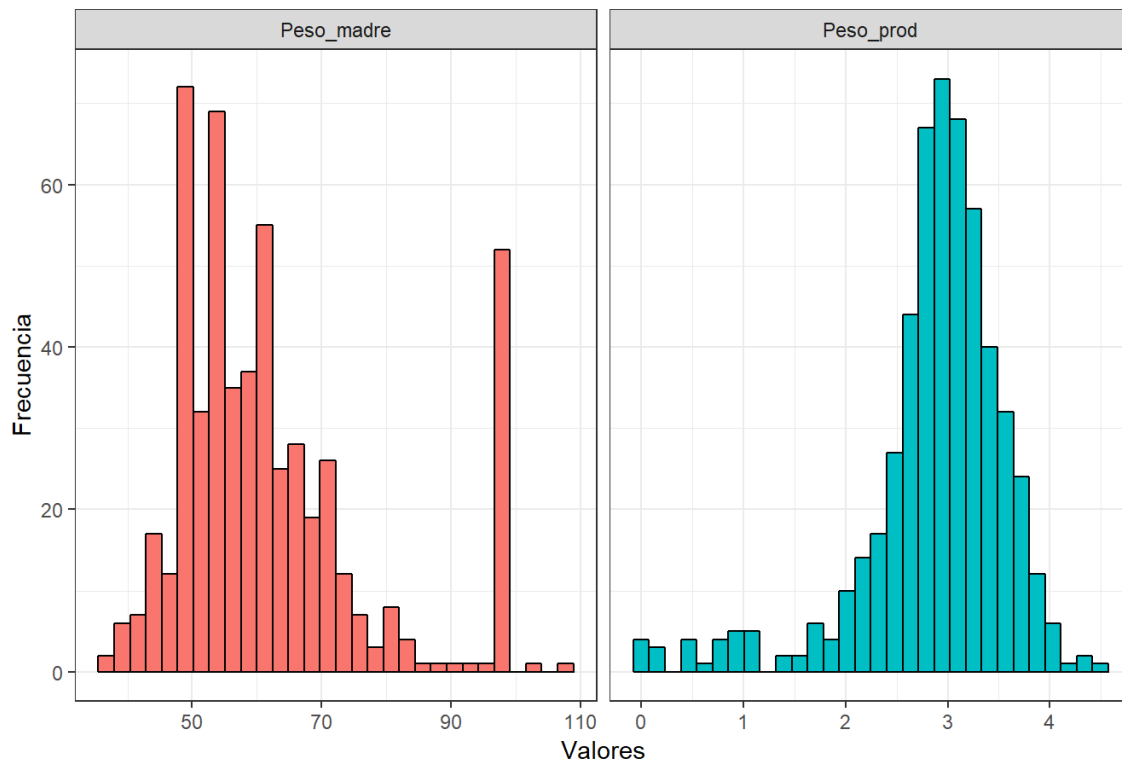
Edad de la madre



El peso del producto tiene una distribución sesgada a la izquierda y la variable se concentra alrededor de los 3 kg. Se observan pesos menores 1 kg. y esto se apoya del resumen de los datos obtenidos previamente.

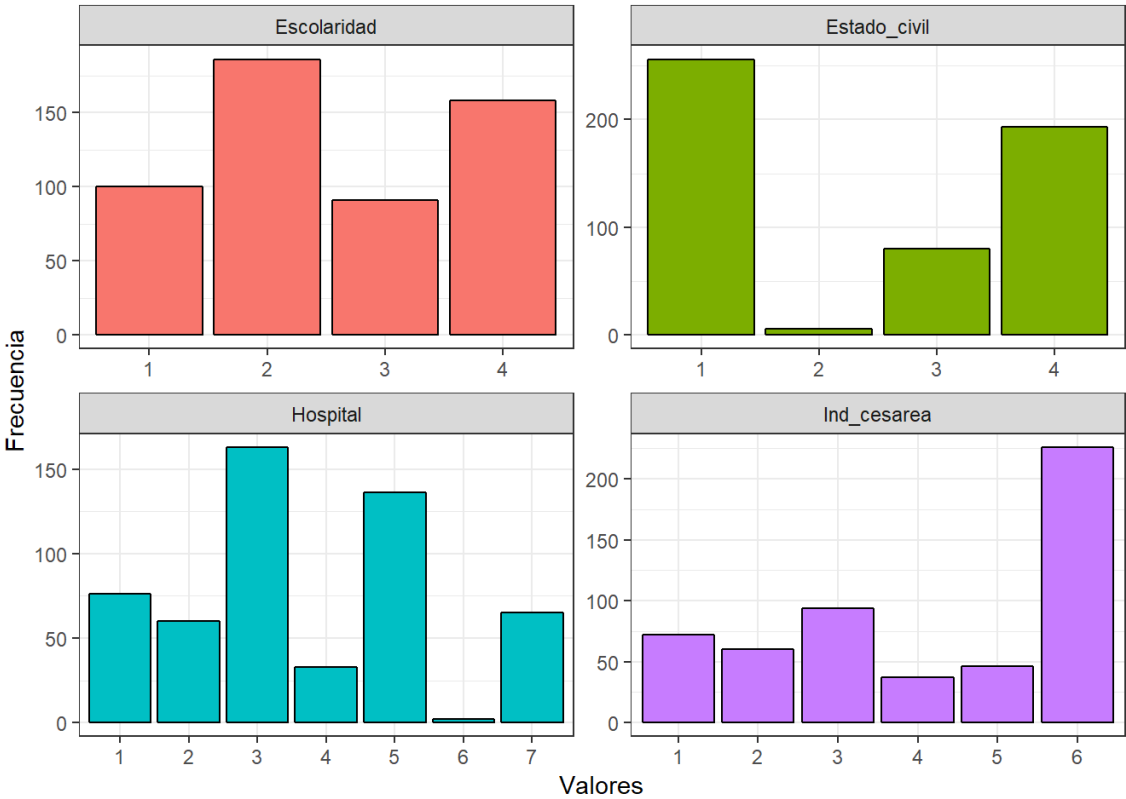
En cuanto al peso de la madre, se observa una distribución bimodal, moda 1 en pesos aproximados de 50 kg y moda 2: pesos aproximados de 100 kg.

Pesos: madre y producto



En el siguiente conjunto de gráficas se pueden observar las frecuencias de las variables categóricas para más de 2 clases (esto para no ser repetitivos con la información mostrada en el resumen).

- La mayoría de las madres tienen nivel de escolaridad secundaria (nivel 2), seguidas de universidad (nivel 4)
- Más de 200 madres son casadas (estado 1) y muy pocas son divorciadas (estado 2)
- Del hospital 6 (Pemex Sur) se tienen muy pocas observaciones y como se menciona anteriormente, en dos hospitales (3 y 5) ocurre la mayor concentración de los nacimientos.
- Poco más de 200 nacimientos fueron vaginales (causa 6). Las causas más comunes en la muestra de la cesárea son: emergencia fetal (causa 3) y DCP (causa 1).



A partir del resumen de los datos y gráficas anteriores, una pregunta que podría resultar de interés es conocer cuáles son las variables que tienen un efecto sobre la mortalidad de los bebés.

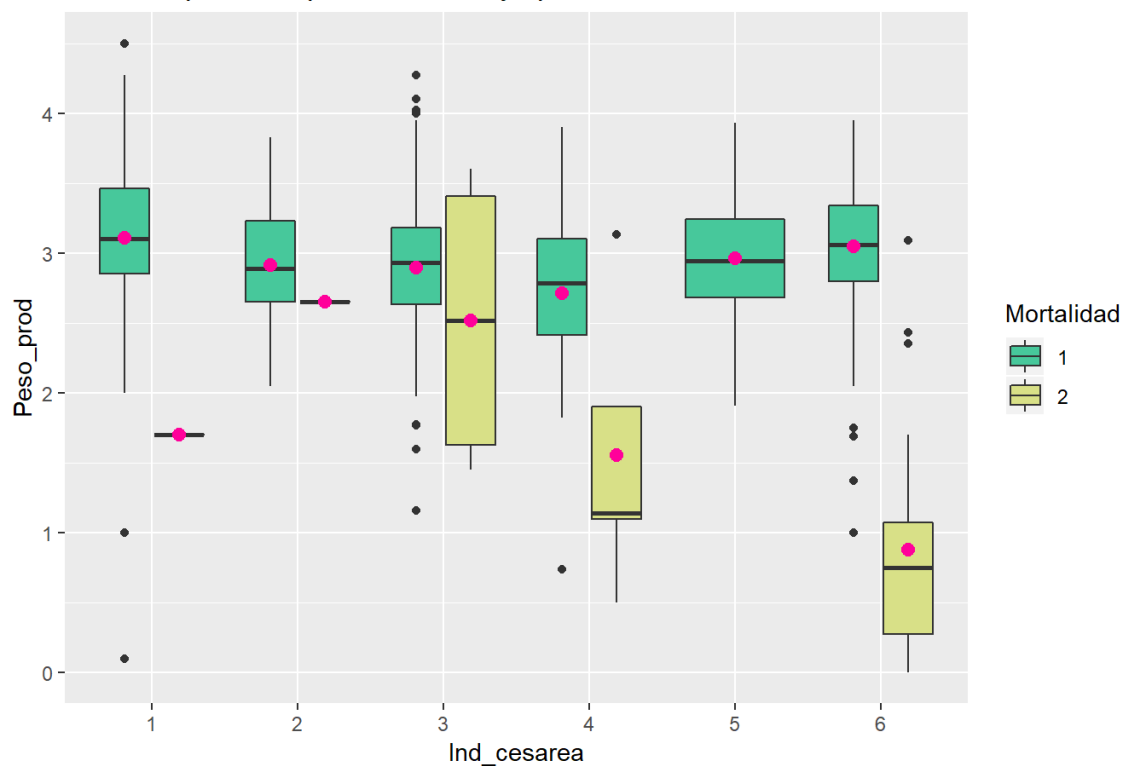
En la tabla, se puede observar que del 100% de los casos de mortalidad, el 62.8571429% ocurrió en partos vaginales.

La única causa que no reporto muertes fue la cesárea por causa iterativa. En relación a su misma causa, la causa 4: Emergencia materna es la que mayor porcentaje de mortalidad tiene.

##						
##		1	2	3	4	5 6
##	1	71	59	88	32	46 204
##	2	1	1	6	5	0 22

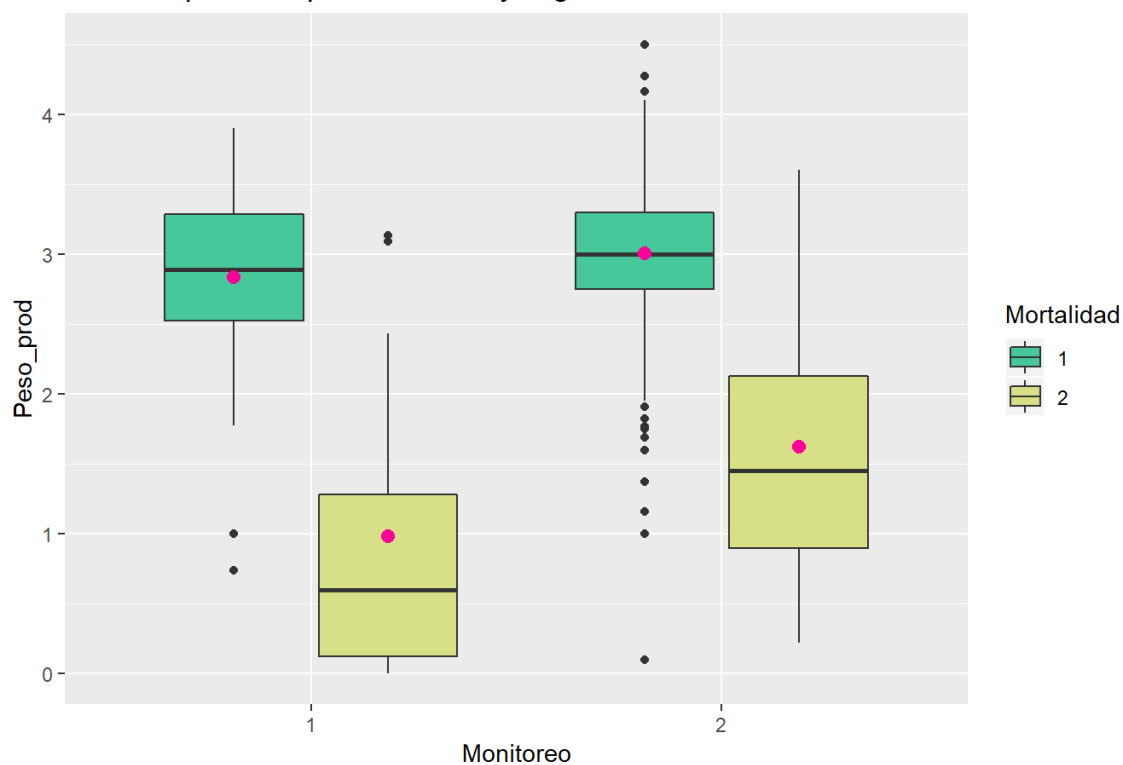
Se observa que en las causas 3 y 4 de las cesáreas y en los partos vaginales (causa 6) hay muertes hay casos de embarazos con muertes. Para la causa 3, el peso mediano de los bebés que murieron es un poco más bajo que el de los bebés vivos. Por el contrario, para la causa 4 y 6 (nacimiento vaginal) la distrubución del peso del producto es muy distinta entre los vivos y los muertos. En rosa, se resalta el peso medio.

Peso del producto por mortalidad y tipo de cesárea



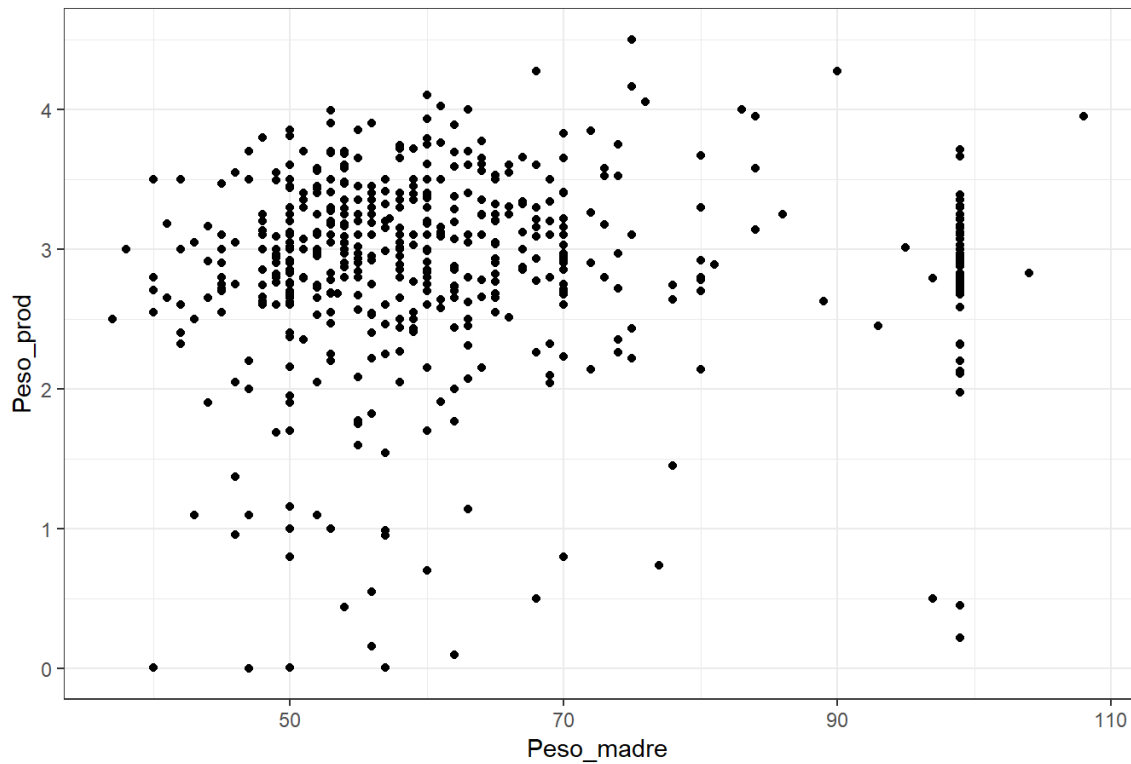
El peso mediano de los bebés vivos que tuvieron monitoreo es relativamente mayor al peso mediando de los vivos que no lo tuvieron. Por el contrario, el peso mediano de los bebés muertos que no tuvieron monitoreo es aproximadamente 0.5 kg mientras que el de los que sí lo tuvieron es de 1.5 kg.

Peso del producto por mortalidad y según monitoreo



En cuanto a la relación peso de la madre y peso del producto, no se observa una relación positiva ni negativa.

Peso de la madre vs hijo



Construcción del modelo

Como se menciona en la sección del análisis exploratorio de los datos, resulta de interés conocer las variables que tienen un efecto en la mortalidad.

Si consideramos la mortalidad como $\{1=\text{muere}, 0=\text{vive}\}$, podemos decir que la variable de interés sigue una distribución Bernoulli. Si quisiéramos considerar a la variable como la tasa de mortalidad, podríamos pensar en una distribución beta para la variable de interés. En este último caso, necesitaríamos calcular el número de muertes y el total de expuestos bajo algún criterio.

Para considerar los datos de manera desagregada, definimos a la variable $Y_i | \pi_i \sim Be(\pi_i)$; es decir,

$$Y_i | \pi_i = \begin{cases} 1 & \text{si el niño nace muerto} \\ 0 & \text{si el niño nace vivo} \end{cases}$$

Transformaciones de las variables

- La variable de *mortalidad* toma valores de 1 y 2 en los datos. Por lo que para poderla modelar con una distribución bernoulli, le restaremos 1.
- Consideramos una variable de interacción entre *indicadora_cesarea* y *forma_nacimiento*

Consideraremos las siguientes ligas:

- liga logística:

$$\pi_i = \frac{e^{X_i' \beta}}{1 + e^{X_i' \beta}}$$

- liga normal:

$$\pi_i = \phi(X_i' \beta)$$

- liga log log complementaria:

$$\pi_i = 1 - e^{-e^{X_i' \beta}}$$

En cualquiera de los casos anteriores, definimos $\eta_i = X_i' \beta$

$$\eta_i = \alpha + \beta_1 * edad_i + \beta_2 * pmadre_i + \beta_3 * pprod_i + \gamma_{z1i} + \delta_{z2i} + \nu_{z3i} + \nu_{z4i} + \nu_{z5i} + \nu_{z6i} + \nu_{z7i} + \nu_{z8i} + \nu_{z5i*z7i}$$

en donde

- z_{1i} uso de anestesia durante parto: $z_{1i} \in \{1, 2\}$
- z_{2i} complicación durante embarazo: $z_{2i} \in \{1, 2\}$
- z_{3i} escolaridad de la madre: $z_{3i} \in \{1, 2, 3, 4\}$
- z_{4i} estado civil de la madre: $z_{4i} \in \{1, 2, 3, 4\}$
- z_{5i} tipo de parto: $z_{5i} \in \{1, 2\}$
- z_{6i} hospital: $z_{6i} \in \{1, \dots, 7\}$
- z_{7i} motivo de la cesárea: $z_{7i} \in \{1, \dots, 6\}$
- z_{8i} monitoreo durante embarazo: $z_{8i} \in \{1, 2, 3, 4\}$

Al tener variables explicativas categóricas, imponemos restricciones de estimabilidad sobre algunos parámetros:

$$\sum_{i=1}^2 \gamma_i = \sum_{i=1}^2 \delta_i = \sum_{i=1}^4 \nu_i = \sum_{i=1}^4 \nu_{1i} = \sum_{i=1}^2 \nu_{2i} = \sum_{i=1}^7 \nu_{3i} = \sum_{i=1}^6 \nu_{4i} = \sum_{i=1}^2 \nu_{5i} = \sum_{i=1}^2 \sum_{j=1}^6 \nu_{6ij} = 0$$

Ejecución del modelo

Se ejecutarán 3 modelos, con diferentes ligas y nos quedaremos con el modelo con el menor DIC. Es importante mencionar que los modelos tienen un periodo de calentamiento de tamaño 3,000 y el número de iteraciones es de 30,000. Asimismo, se verificó que cada uno de los parámetros converge, y para reducir papel, no se muestran aquí las gráficas del monitoreo de la cadena.

En todos los modelos, el intercepto se puede interpretar como el nivel medio global ya que los datos numéricos están centrados.

- Coeficientes de las variables continuas para el modelo logístico.

```
##           mean      2.5%    97.5%
## beta1 -0.04958036 -0.17380  0.06425
## beta2 -0.02016765 -0.06852  0.02465
## beta3 -3.42847200 -4.60100 -2.44100
```

La única variable significativa es el peso del producto. Las otras dos variables, edad y peso de la madre, no son significativas ya que el intervalo de confianza contiene al cero.

- Coeficientes de las variables continuas para el modelo con liga normal

Al igual que en el caso con liga logística, la única variable que resulta significativa es el peso del producto.

```
##           mean      2.5%    97.5%
## beta1 -0.01891853 -0.07183000  0.03350
## beta2 -0.01094753 -0.03439025  0.01058
## beta3 -1.59527448 -2.04200000 -1.18800
```

- Coeficientes de las variables continuas para el modelo con liga *clog log*

Al igual que en el caso con liga normal, la única variable que resulta significativa es el peso del producto.

```
##           mean      2.5%    97.5%
## beta1 -0.004688326 -0.09891000  0.08801
## beta2 -0.012191194 -0.05250025  0.02534
## beta3 -2.751275444 -3.70200000 -1.95200
```

Para las tres ligas, la única variable significativa es el peso del producto. Las otras dos variables, edad y peso de la madres, no son significativas ya que el intervalo de confianza contiene al cero. Por esta razón, interpretaremos el coeficiente de β_3 para todas las ligas. Únicamente para el mejor modelo, realizaremos el análisis completo.

Si se incrementa en un un kilogramo el peso del producto,

- bajo la liga logístista, el efecto estaría dado por $e(\beta_3)$. En este caso, 0.0324365, es decir, ante un incremento de un kilogramo en el peso del producto, la razón de momios disminuye un 97%.
- bajo la liga normal, el efecto está dado por -1.5952745; es decir, ante un incremento de una unidad, la probabilidad de muerte disminuye.
- bajo la liga complementaria log log, $\beta_3 = -2.7512754$ nos da el cambio en la probabilidad (una vez transformada la variable) ante un incremento de una unidad. En este caso, se tiene que el “momio” de morir es 0.0638464.

Se elegirá el modelo con el menor DIC y a partir de éste, se realizará el análisis de las variables.

liga	DIC
logística	127.3
normal	127.3
clog-log	124.7
log-log	131.5

A partir de la tabla anterior, se puede concluir que el mejor modelo es el que utiliza la liga complementaria log-log.

Para la liga clog-log definimos el “momio” como $\frac{-\log(1-\pi_j)}{-\log(1-\pi_i)}$.

Recordemos que para las variables numéricas, únicamente la del peso del producto resultó significativa.

A continuación se muestran las gráficas de las variables categóricas para determinar los efectos:

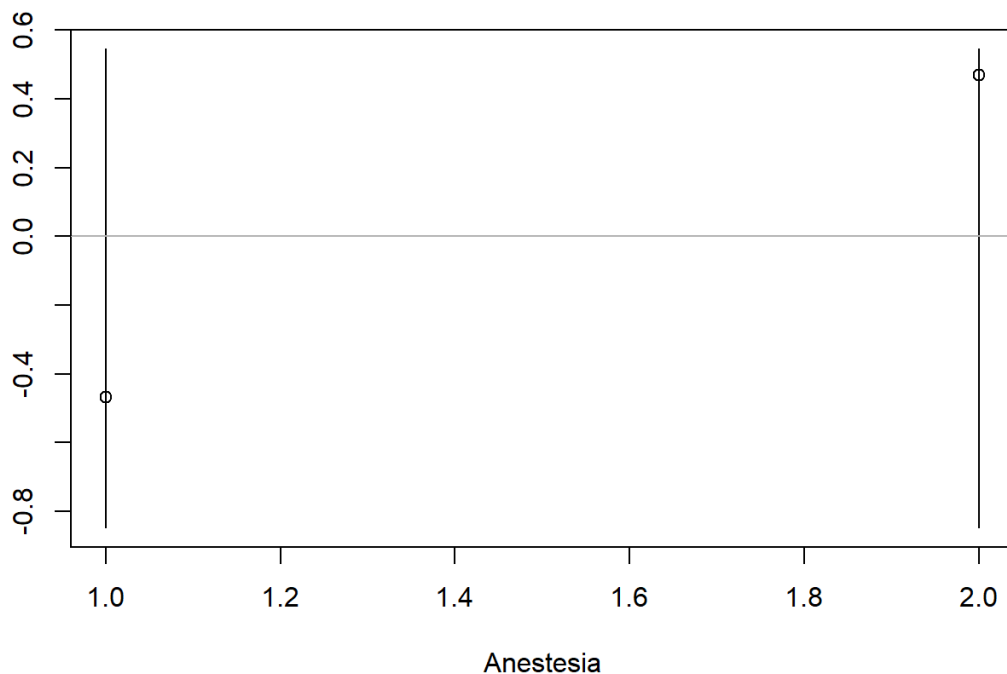
(nota: las variable son categóricas, pero no logré hacer que se vieran únicamente los valores de 1 y 2 en el eje de las x).

- Efecto antesia:

Se observa que la variable Anestesia no resulta significativa.

```
## [1] -0.4691840  0.5448168 -0.8477000
```

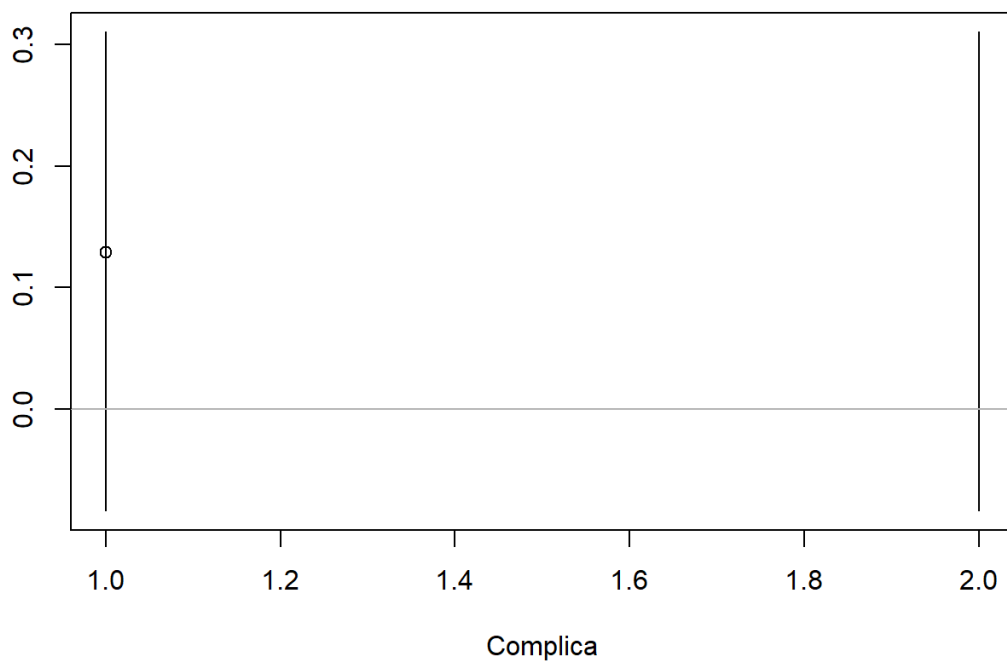

Mortalidad: Efecto de anestesia



- Efecto para la variable complica

La variable no resulta significativa para ninguno de los niveles.

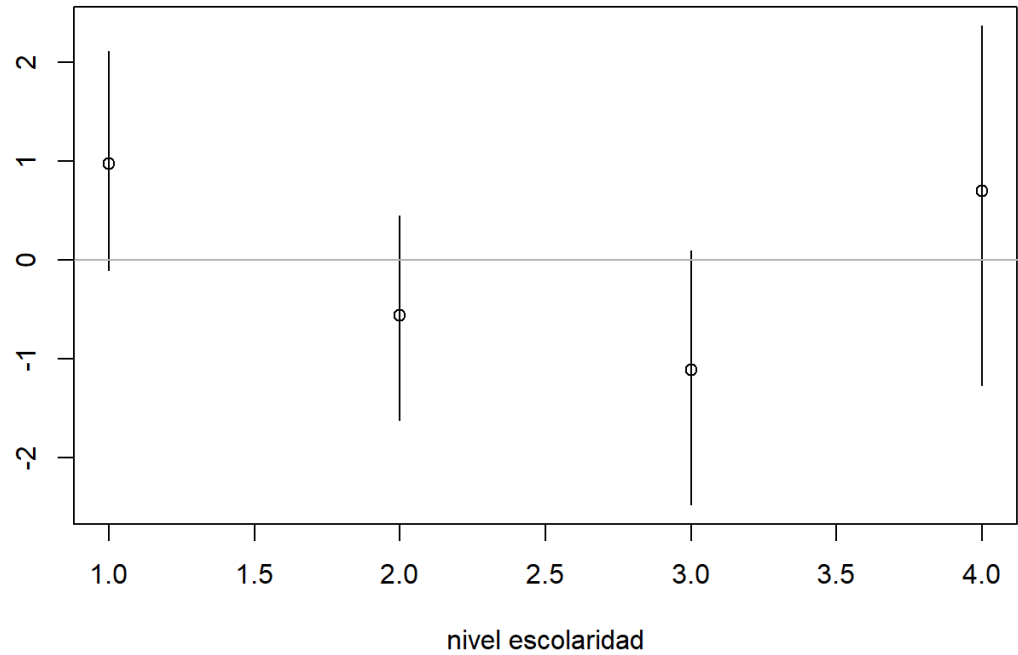
Mortalidad: Efecto complica



- Efecto para la variable escolaridad

En los madres cuyos niveles de escolaridad 2 y 3 (secundaria y preparatoria), el "momio" probabilidad de muerte del producto es menor. Sin embargo, la variable no es significativa para ningún nivel.

Mortalidad: Efecto escolaridad

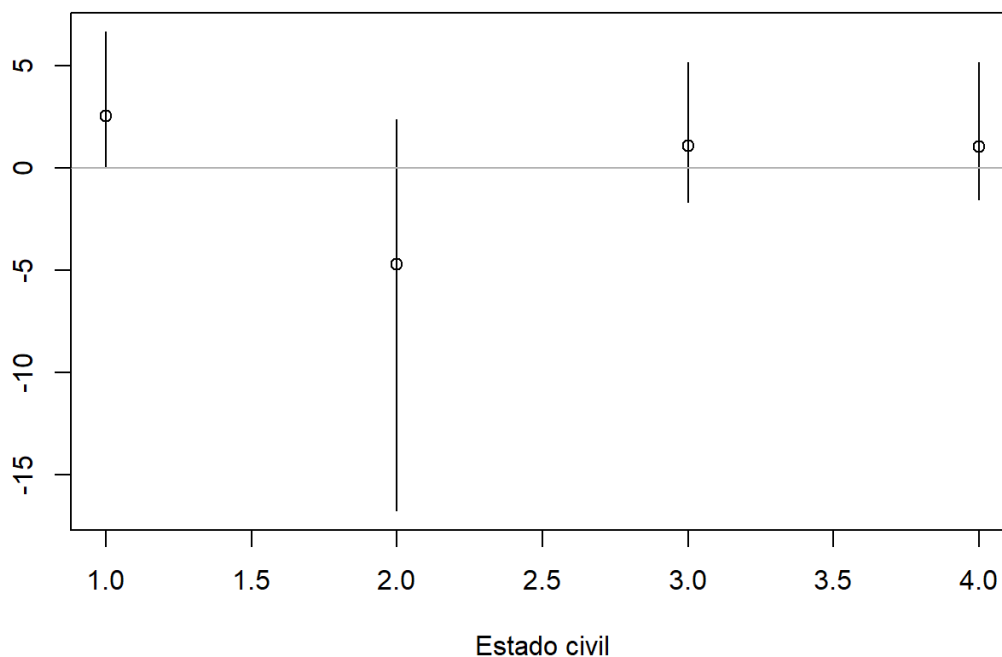


- Efecto para la variable estado civil

El único estado civil que resulta significativo es el 1: casada. Una vez trasformada la probabilidad, nos indica el cambio en el “momio” de muerte.

##		mean	2.5%	97.5%
##	nu1.est[1]	2.555048	0.0275115	6.653000
##	nu1.est[2]	-4.704741	-16.7400000	2.399025
##	nu1.est[3]	1.086460	-1.6630250	5.178075
##	nu1.est[4]	1.063224	-1.5300250	5.181025

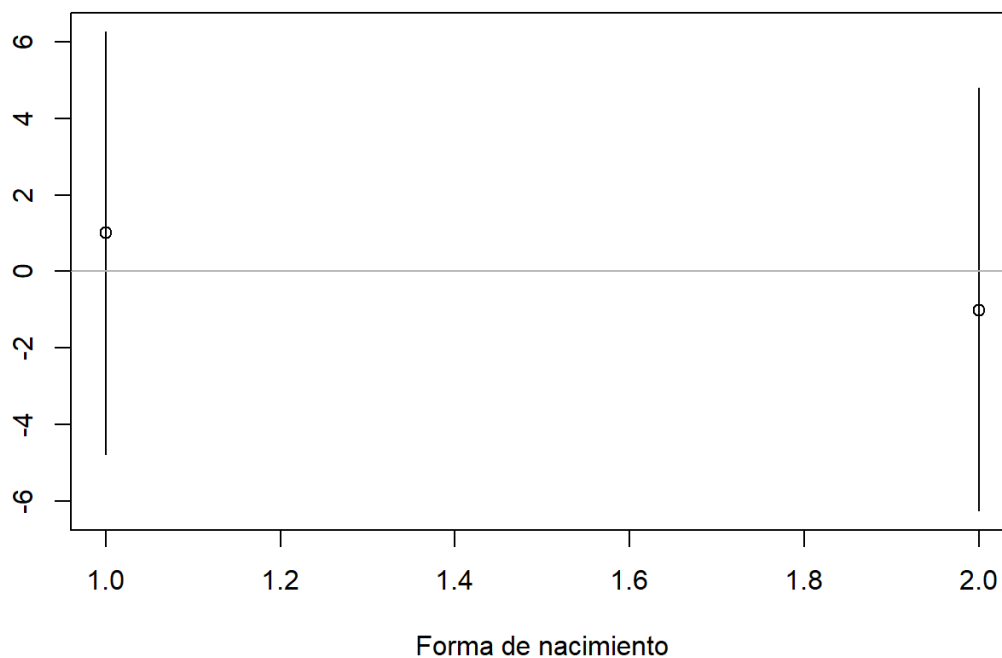
Mortalidad: efecto del estado civil



- Efecto para la forma de nacimiento

La forma de nacimiento no es significativa; se observa que en los partos vaginales la probabilidad de muerte es menor.

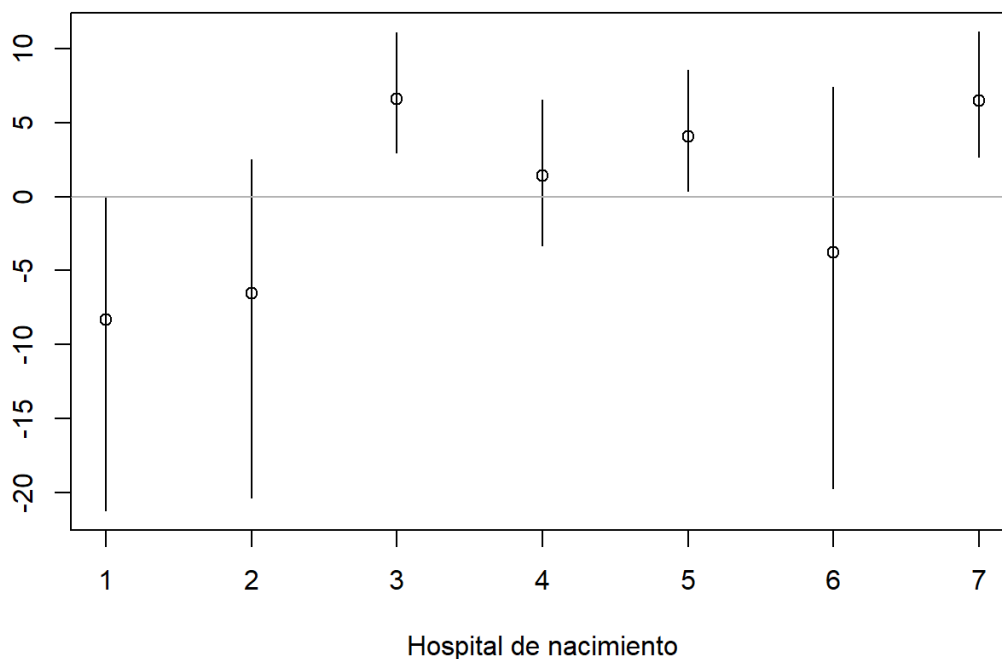
Mortalidad: efecto de la forma de nacimiento



- Efecto para la variable categórica hospital

En el hospital 1 es en donde el “momio” de morir disminuye más. El efecto de los hospitales que resultan significativos son 1,3, 5 y 7: Hospital General, Oaxaca y Tlaxcala. Los hospitales 3 y 7 tienen relativamente los mismos niveles para el “momio”.

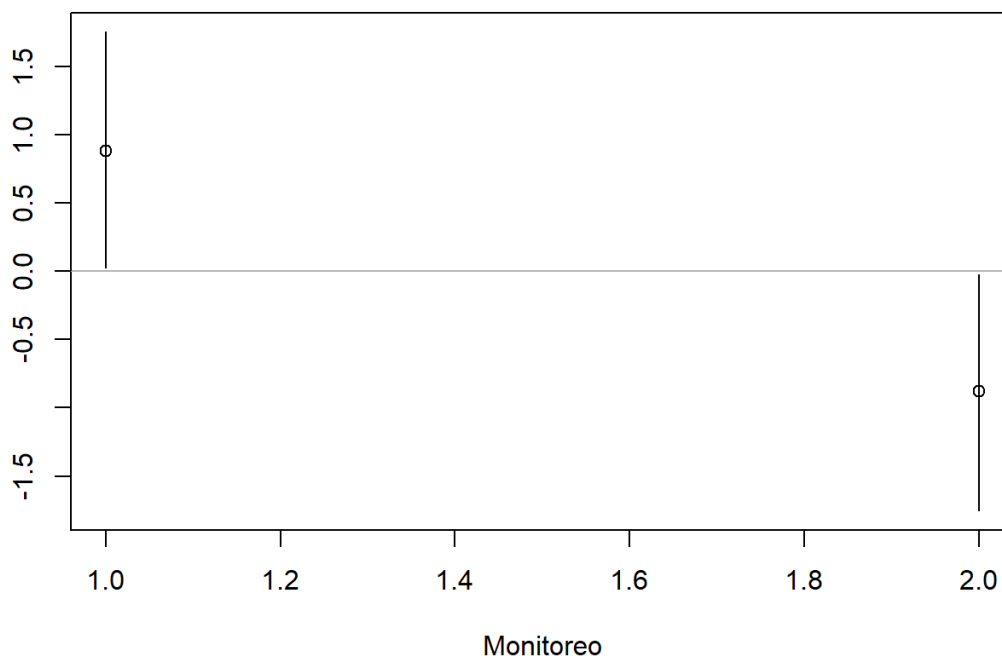
Mortalidad: efecto del hospital



- Efecto de la variable monitoreo

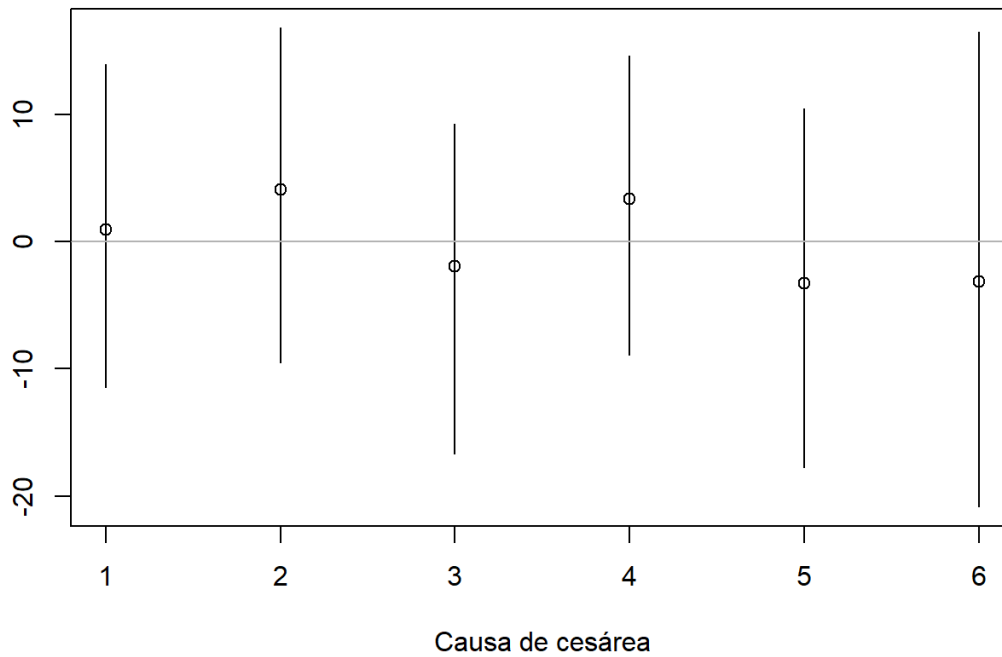
Ambos efectos (no monitoreo y monitoreo), resultan significativos. El “momio” de no monitorear a la madre durante el embarazo es mucho más grande que el “momio” de monitorearla.

Mortalidad: efecto del monitoreo



- Efecto de la variable causa de cesárea

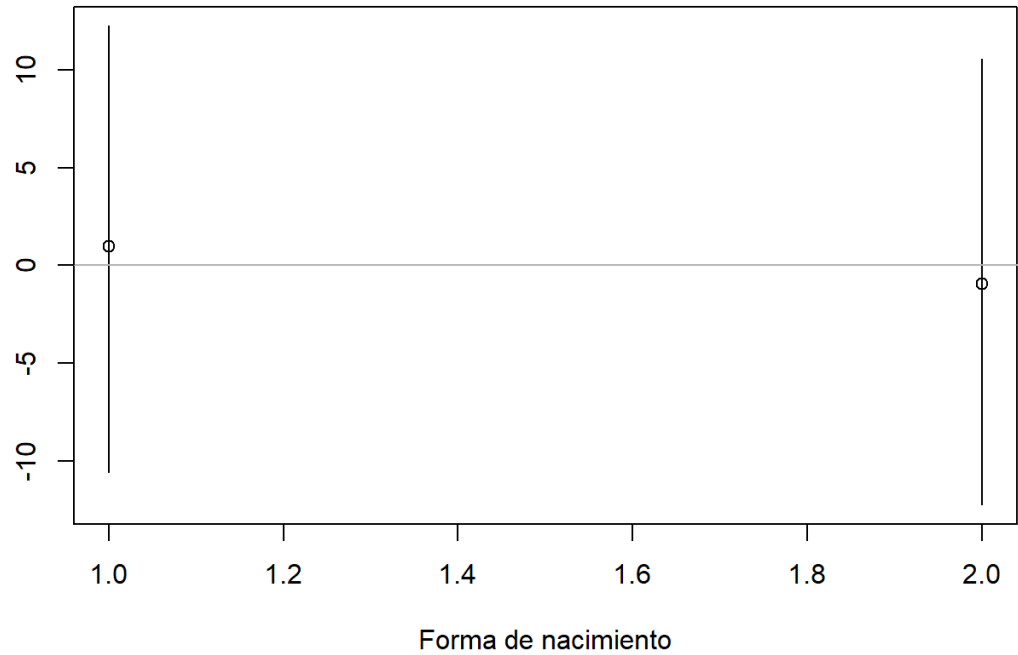
La variable causa de la cesárea no resulta significativa para ninguna causa.

Mortalidad: efecto de la causa de cesárea

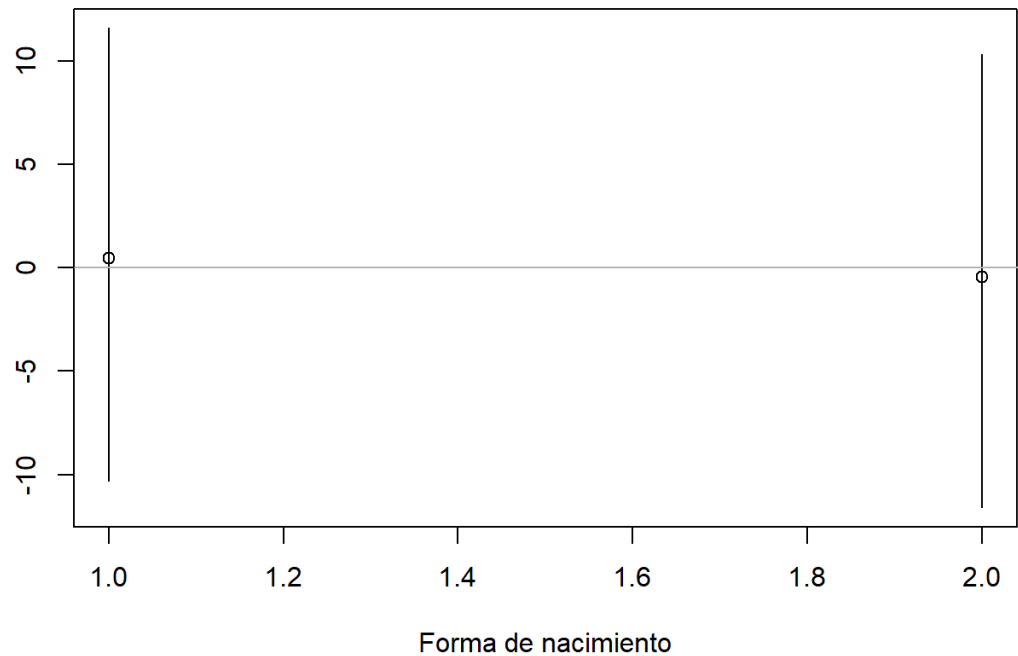
- Efecto para la variable de interacción: forma de nacimiento y motivo de la cesárea

La variable de interacción no resultó significativa para ninguna causa

DCP

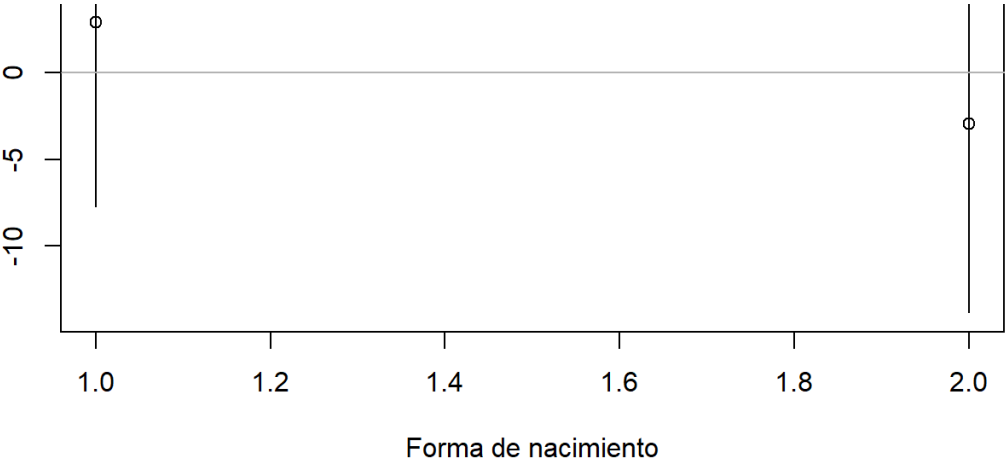


Electiva

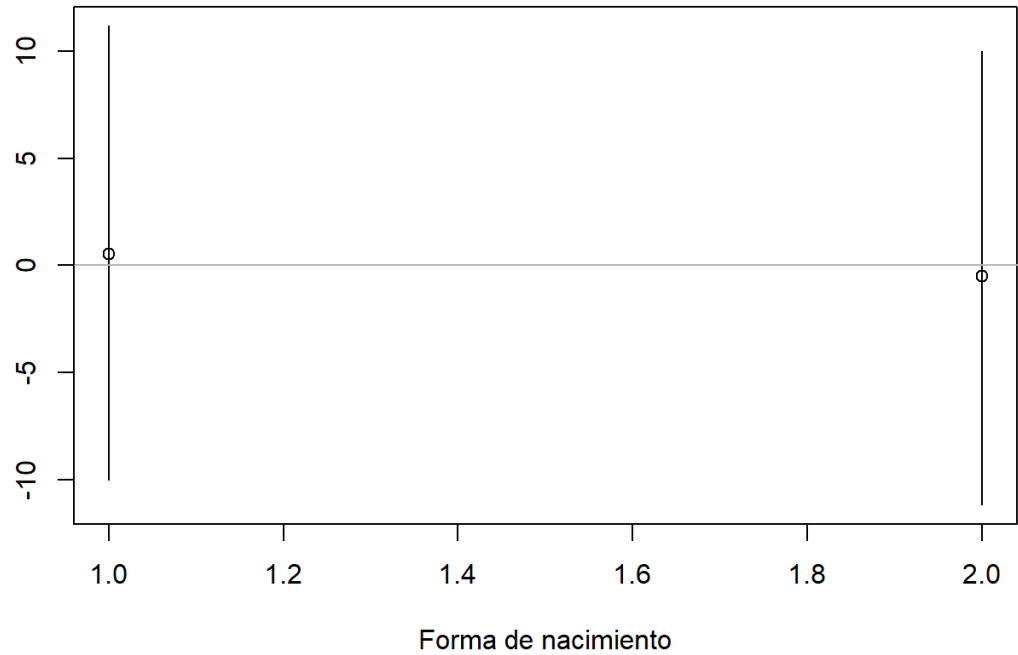


Emer. Fetal

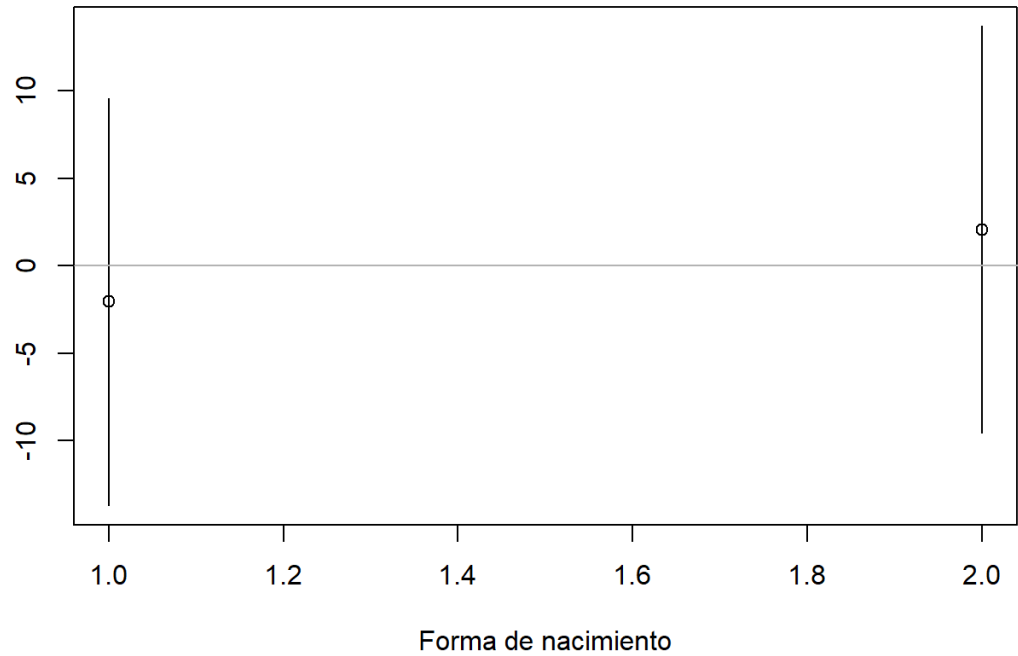




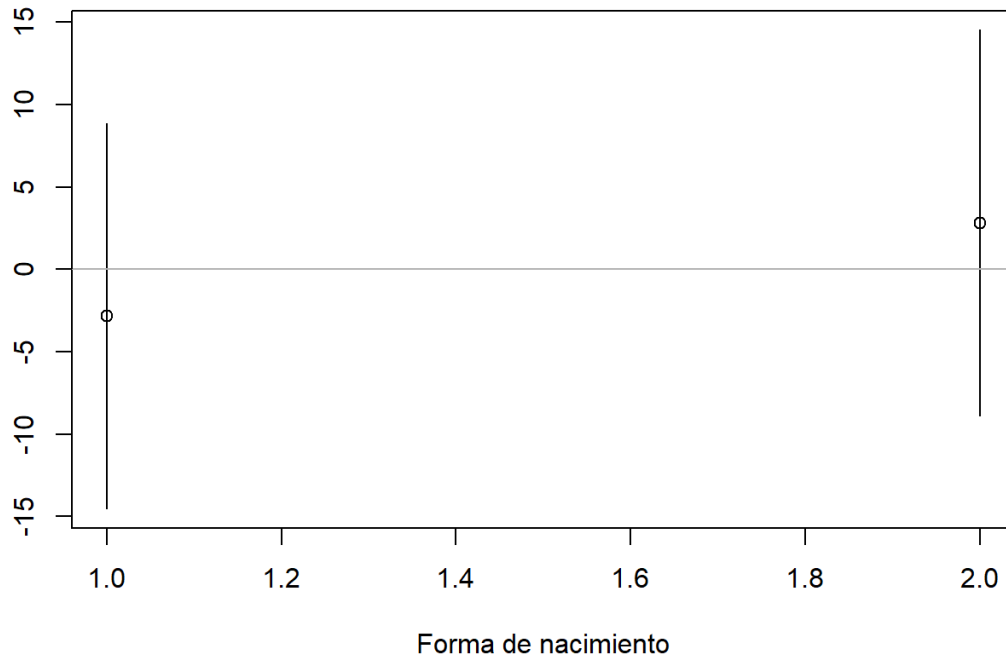
Emer. Mat



Iterativa



No hubo



Código

En la primera sección se detalla el código Bugs y en la segunda, el código en R para su ejecución.

- Liga Logística

```

modelString=
  "model{
    for (i in 1:n) {
      y[i] ~ dbern(p[i])
      mu[i] <- p[i]
      eta[i] <- alpha + beta1*edad[i] + beta2*pmadre[i]+ beta3*pprod[i] +
      gama[anestesia[i]] + delta[complica[i]] + nu[escolaridad[i]] + nu1[estado[i]]+
      nu2[forma[i]]+nu3[hospital[i]] + nu5[monitoreo[i]] + nu4[forma[i], cesarea[i]]
      p[i] <- exp(eta[i])/(1+exp(eta[i]))
    }

    # Priors
    alpha ~ dnorm(0, 0.001)
    beta1 ~ dnorm(0, 0.001)
    beta2 ~ dnorm(0, 0.001)
    beta3 ~ dnorm(0, 0.001)
    for (i in 1:2) {gama[i] ~ dnorm(0, 0.01)}
    for (i in 1:2) {delta[i] ~ dnorm(0, 0.01)}
    for (i in 1:4) {nu[i] ~ dnorm(0, 0.01)}
    for (i in 1:4) {nu1[i] ~ dnorm(0, 0.01)}
    for (i in 1:2) {nu2[i] ~ dnorm(0, 0.01)}
    for (i in 1:7) {nu3[i] ~ dnorm(0, 0.01)}
    for (i in 1:2) {for (j in 1:6) {
      nu4[i,j] ~ dnorm(0,0.01)
    }}
    for (i in 1:2) {nu5[i] ~ dnorm(0, 0.01)}

    # restricciones de estimabilidad
    alpha.est <- alpha + mean(gama[])+mean(delta[])+mean(nu[])+mean(nu1[])+mean(nu2[])+mean(nu3[])+mean(nu5[])+mean(nu
    4[,])
    for (i in 1:2) {gama.est[i] <- gama[i] - mean(gama[]) }
    for (i in 1:2) {delta.est[i] <- delta[i] - mean(delta[])}
    for (i in 1:4) {nu.est[i] <- nu[i] - mean(nu[])}
    for (i in 1:4) {nu1.est[i] <- nu1[i] - mean(nu1[]) }
    for (i in 1:2) {nu2.est[i] <- nu2[i] - mean(nu2[])}
    for (i in 1:7) {nu3.est[i] <- nu3[i] - mean(nu3[])}
    for (i in 1:2) { for(j in 1:6) {
      nu4.est[i,j] <- nu4[i,j] - mean(nu4[i, ]) - mean(nu4[,j]) + mean(nu4[,]) }
    }

    for (i in 1:2) {nu5.est[i] <- nu5[i] - mean(nu5[])}

    # Prediction1
    for (i in 1:n) {
      yf1[i] ~ dbern(p[i])
    }
  }
"
writeLines(modelString, con="modelo_gestacional.txt")

```

- Liga normal

```

modelString="
model{
  for (i in 1:n) {
    y[i] ~ dbern(p[i])
    mu[i] <- p[i]
    eta[i] <- alpha + beta1*edad[i] + beta2*pmadre[i]+ beta3*pprod[i] +
    gama[anestesia[i]] + delta[complica[i]] + nu[escolaridad[i]] + nu1[estado[i]]+
    nu2[forma[i]]+nu3[hospital[i]] + nu5[monitoreo[i]] + nu4[forma[i], cesarea[i]] + nu6[cesarea[i]]
    p[i] <- phi(eta[i])
  }

# Priors
alpha ~ dnorm(0, 0.001)
beta1 ~ dnorm(0, 0.001)
beta2 ~ dnorm(0, 0.001)
beta3 ~ dnorm(0, 0.001)
for (i in 1:2) {gama[i] ~ dnorm(0, 0.01)}
for (i in 1:2) {delta[i] ~ dnorm(0, 0.01)}
for (i in 1:4) {nu[i] ~ dnorm(0, 0.01)}
for (i in 1:4) {nu1[i] ~ dnorm(0, 0.01)}
for (i in 1:2) {nu2[i] ~ dnorm(0, 0.01)}
for (i in 1:7) {nu3[i] ~ dnorm(0, 0.01)}
for (i in 1:2) {for (j in 1:6) {
  nu4[i,j] ~ dnorm(0,0.01)
}}
for (i in 1:2) {nu5[i] ~ dnorm(0, 0.01)}
for (i in 1:6) {nu6[i] ~ dnorm(0, 0.01)}

# restricciones de estimabilidad
alpha.est <- alpha + mean(gama[])+mean(delta[])+mean(nu[])+mean(nu1[])+mean(nu2[])+mean(nu3[])+mean(nu5[])+mean(nu
4[,]) + mean(nu6[])
for (i in 1:2) {gama.est[i] <- gama[i] - mean(gama[])}}
for (i in 1:2) {delta.est[i] <- delta[i] - mean(delta[])}}
for (i in 1:4) {nu.est[i] <- nu[i] - mean(nu[])}}
for (i in 1:4) {nu1.est[i] <- nu1[i] - mean(nu1[])}}
for (i in 1:2) {nu2.est[i] <- nu2[i] - mean(nu2[])}}
for (i in 1:7) {nu3.est[i] <- nu3[i] - mean(nu3[])}}
for (i in 1:2) { for(j in 1:6) {
  nu4.est[i,j] <- nu4[i,j] - mean(nu4[i, ]) - mean(nu4[,j]) + mean(nu4[,])}
}

for (i in 1:2) {nu5.est[i] <- nu5[i] - mean(nu5[])}}
for (i in 1:6) {nu6.est[i] <- nu6[i] - mean(nu6[])}}

# Prediction1
for (i in 1:n) {
  yf1[i] ~ dbern(p[i])
}
}
"
writeLines(modelString, con="modelo_gestacional_probit.txt")

```

- Liga clog log

```

modelString=
  "model{
    for (i in 1:n) {
      y[i] ~ dbern(p[i])
      mu[i] <- p[i]
      eta[i] <- alpha + beta1*edad[i] + beta2*pmadre[i]+ beta3*pprod[i] +
      gama[anestesia[i]] + delta[complica[i]] + nu[escolaridad[i]] + nu1[estado[i]]+
      nu2[forma[i]]+nu3[hospital[i]] + nu5[monitoreo[i]] + nu4[forma[i], cesarea[i]] + nu6[cesarea[i]]
      cloglog(p[i]) <- eta[i]
    }

# Priors
alpha ~ dnorm(0, 0.001)
beta1 ~ dnorm(0, 0.001)
beta2 ~ dnorm(0, 0.001)
beta3 ~ dnorm(0, 0.001)
for (i in 1:2) {gama[i] ~ dnorm(0, 0.01)}
for (i in 1:2) {delta[i] ~ dnorm(0, 0.01)}
for (i in 1:4) {nu[i] ~ dnorm(0, 0.01)}
for (i in 1:4) {nu1[i] ~ dnorm(0, 0.01)}
for (i in 1:2) {nu2[i] ~ dnorm(0, 0.01)}
for (i in 1:7) {nu3[i] ~ dnorm(0, 0.01)}
for (i in 1:2) {for (j in 1:6) {
  nu4[i,j] ~ dnorm(0,0.01)
}}
for (i in 1:2) {nu5[i] ~ dnorm(0, 0.01)}
for (i in 1:6) {nu6[i] ~ dnorm(0, 0.01)}

# restricciones de estimabilidad
alpha.est <- alpha + mean(gama[])+mean(delta[])+mean(nu[])+mean(nu1[])+mean(nu2[])+mean(nu3[])+mean(nu5[])+mean(nu
4[,]) + mean(nu6[])
for (i in 1:2) {gama.est[i] <- gama[i] - mean(gama[])}}
for (i in 1:2) {delta.est[i] <- delta[i] - mean(delta[])}}
for (i in 1:4) {nu.est[i] <- nu[i] - mean(nu[])}}
for (i in 1:4) {nu1.est[i] <- nu1[i] - mean(nu1[])}}
for (i in 1:2) {nu2.est[i] <- nu2[i] - mean(nu2[])}}
for (i in 1:7) {nu3.est[i] <- nu3[i] - mean(nu3[])}}
for (i in 1:2) { for(j in 1:6) {
  nu4.est[i,j] <- nu4[i,j] - mean(nu4[i, ]) - mean(nu4[,j]) + mean(nu4[,])}
}

for (i in 1:2) {nu5.est[i] <- nu5[i] - mean(nu5[])}}
for (i in 1:6) {nu6.est[i] <- nu6[i] - mean(nu6[])}}

# Prediction1
for (i in 1:n) {
  yf1[i] ~ dbern(p[i])
}
}
"

writeLines(modelString, con="modelo_gestacional_cloglog.txt")

```

Código de R

```

# directorio y carga de Librerías
setwd("~/MCD/Otoño2019/Regresión Avanzada/Examen parcial/")
library(R2OpenBUGS)
library(R2jags)
library(ggplot2)
library(dplyr)
library(tidyr)

#Lectura de datos
gestacional <- read.csv("Gestacional.csv", header=TRUE)
head(gestacional)
gestacional1 <- gestacional

gestacional1$Anestesia <- as.factor(gestacional1$Anestesia)
gestacional1$Complica <- as.factor(gestacional1$Complica)
gestacional1$Escolaridad <- as.factor(gestacional1$Escolaridad)
gestacional1$Estado_civil <- as.factor(gestacional1$Estado_civil)
gestacional1$Forma_nac <- as.factor(gestacional1$Forma_nac)
gestacional1$Hospital <- as.factor(gestacional1$Hospital)
gestacional1$Ind_cesarea <- as.factor(gestacional1$Ind_cesarea)
gestacional1$Monitoreo <- as.factor(gestacional1$Monitoreo)
gestacional1$Mortalidad <- as.factor(gestacional1$Mortalidad)
summary(gestacional1)

gestacional %>%
  gather(Attributes, value, c(3)) %>%
  ggplot(aes(x=value, fill=Attributes)) +
  geom_histogram(colour="black", show.legend=FALSE, binwidth = 2) +
  facet_wrap(~Attributes, scales="free") +
  labs(x="Valores", y="Frecuencia", title="Edad de la madre") +
  theme_bw()

gestacional %>%
  gather(Attributes, value, c(11, 12)) %>%
  ggplot(aes(x=value, fill=Attributes)) +
  geom_histogram(colour="black", show.legend=FALSE) +
  facet_wrap(~Attributes, scales="free_x") +
  labs(x="Valores", y="Frecuencia", title="Pesos: madre y producto") +
  theme_bw()

gestacional1 %>%
  gather(Attributes, count, c(4,5, 7,8)) %>%
  ggplot(aes(x=count, fill=Attributes)) +
  geom_bar(colour="black", show.legend=FALSE) +
  facet_wrap(~Attributes, scales="free") +
  labs(x="Valores", y="Frecuencia") +
  theme_bw()

table(gestacional1$Mortalidad, gestacional$Ind_cesarea)

ggplot(gestacional1, aes(x=Ind_cesarea, y=Peso_prod, fill=Mortalidad)) +
  geom_boxplot() +
  ggtitle("Peso del producto por mortalidad y tipo de cesárea")+
  scale_fill_manual(values=c("#47c89b", "#d8e087"))+
  stat_summary(fun.y = mean, color = "#ff009a", position = position_dodge(0.75),
    geom = "point", shape = 20, size = 4,
    show.legend = FALSE)

ggplot(gestacional1, aes(x=Monitoreo, y=Peso_prod, fill=Mortalidad)) +
  geom_boxplot() +
  ggtitle("Peso del producto por mortalidad y según monitoreo")+
  scale_fill_manual(values=c("#47c89b", "#d8e087"))+

```

```

stat_summary(fun.y = mean, color = "#ff009a", position = position_dodge(0.75),
             geom = "point", shape = 20, size = 4,
             show.legend = FALSE)

ggplot(gestacional, aes(x=Peso_madre, y=Peso_prod)) +
  geom_point() +
  labs(title="Peso de la madre vs hijo") +
  theme_bw()

#### creación conjuntos para ejecución de modelo
n<-dim(gestacional)[1]
data<-list("n"=n,"y"=gestacional$Mortalidad-1, "edad"=gestacional$Edad-mean(gestacional$Edad), "pmadre"=gestacional$Peso_madre-mean(gestacional$Peso_madre), "pprod"=gestacional$Peso_prod-mean(gestacional$Peso_prod), "anestesia"=gestacional$Anestesia, "complica"=gestacional$Complica, "escolaridad"=gestacional$Escolaridad, "estado"=gestacional$Estado_civil, "forma"=gestacional$Forma_nac, "hospital"=gestacional$Hospital, "cesarea"=gestacional$Ind_cesarea, "monitoreo"=gestacional$Monitoreo)

#-Definiendo inits-
inits<-function(){list(alpha=0, beta1=0, beta2=0, beta3=0,gama=rep(0,2),delta=rep(0,2), nu=rep(0,4), nu1=rep(0,4), nu2=rep(0,2), nu3=rep(0,7), nu5=rep(0,2),yf1=rep(1,n), nu4=matrix(rep(0,12), ncol=6), nu6=rep(0,6))}

parameters <- c("alpha.est","beta1", "beta2", "beta3","gama.est","delta.est","nu.est","nu1.est", "nu2.est", "nu3.est", "nu4.est", "nu5.est", "nu6.est", "yf1")

mod1.sim<- bugs(data,inits,parameters,model.file="modelo_gestacional.txt",
               n.iter=30000,n.chains=2,n.burnin=3000,n.thin=1, debug=TRUE)

mod1a.sim<- bugs(data,inits,parameters,model.file="modelo_gestacional_probit.txt",
                 n.iter=30000,n.chains=2,n.burnin=3000,n.thin=1, debug=TRUE)

mod1b.sim<- bugs(data,inits,parameters,model.file="modelo_gestacional_cloglog.txt",
                 n.iter=30000,n.chains=2,n.burnin=3000,n.thin=1, debug=TRUE)

out.sum<-mod1.sim$summary
out.sum1a<-mod1a.sim$summary
out.sum1b<-mod1b.sim$summary

#Tabla resumen
out.sum.t<-out.sum[grep("beta",rownames(out.sum)),c(1,3,7)] #logistico
out.sum.t1a<-out.sum1a[grep("beta",rownames(out.sum1a)),c(1,3,7)] #probit
out.sum.t1b<-out.sum1b[grep("beta",rownames(out.sum1b)),c(1,3,7)] # cloglog

print(out.sum.t)
print(out.sum.t1a)
print(out.sum.t1b)
#DIC
mod1.sim$DIC
mod1a.sim$DIC
mod1b.sim$DIC
mod1c.sim$DIC

#Monitoreo de las cadenas
traceplot(mod1.sim)
traceplot(mod1a.sim)
traceplot(mod1b.sim)
traceplot(mod1c.sim)

#### Coeficientes del modelo
#gama: anestesia
out.gama<-out.sum1b[grep("gama",rownames(out.sum1b)),]
out.est<-out.gama

```

```

k<-2
a <- as.factor(c(1,2))
print(out.est[c(1,3,7)])
ymin<-min(out.est[c(1,3,7)])
ymax<-max(out.est[c(1,3,7)])
plot(1:k,out.est[,1],xlab="Anestesia",ylab="",ylim=c(ymin,ymax))
segments(1:k,out.est[3],1:k,out.est[7])
abline(h=0,col="grey70")
title("Mortalidad: Efecto de anestesia")

#delta
out.delta<-out.sum1b[grep("delta",rownames(out.sum1b)),]
out.est<-out.delta
k<-2
ymin<-min(out.est[c(1,3,7)])
ymax<-max(out.est[c(1,3,7)])
plot(1:k,out.est[,1],xlab="Complica",ylab="",ylim=c(ymin,ymax))
segments(1:k,out.est[3],1:k,out.est[7])
abline(h=0,col="grey70")
title("Mortalidad: Efecto complica")

#nu
out.nu<-out.sum1b[grep("nu",rownames(out.sum1b)),]
out.est<-out.nu[1:4, ]
k<-4
ymin<-min(out.est[,c(1,3,7)])
ymax<-max(out.est[,c(1,3,7)])
plot(1:k,out.est[,1],xlab="nivel escolaridad",ylab="",ylim=c(ymin,ymax))
segments(1:k,out.est[,3],1:k,out.est[,7])
abline(h=0,col="grey70")
title("Mortalidad: Efecto escolaridad")

#nu1
out.nu1<-out.sum1b[grep("nu1",rownames(out.sum1b)),]
out.est<-out.nu1
k<-4
print(out.est[,c(1,3,7)])
ymin<-min(out.est[,c(1,3,7)])
ymax<-max(out.est[,c(1,3,7)])
plot(1:k,out.est[,1],xlab="Estado civil",ylab="",ylim=c(ymin,ymax))
segments(1:k,out.est[,3],1:k,out.est[,7])
abline(h=0,col="grey70")
title("Mortalidad: efecto del estado civil")

#nu2: forma de nacimiento
out.nu2<-out.sum1b[grep("nu2",rownames(out.sum1b)),]
out.est<-out.nu2
k<-2
ymin<-min(out.est[,c(1,3,7)])
ymax<-max(out.est[,c(1,3,7)])
plot(1:k,out.est[,1],xlab="Forma de nacimiento",ylab="",ylim=c(ymin,ymax))
segments(1:k,out.est[,3],1:k,out.est[,7])
abline(h=0,col="grey70")
title("Mortalidad: efecto de la forma de nacimiento")

#nu3: hospital
out.nu3<-out.sum1b[grep("nu3",rownames(out.sum1b)),]
out.est<-out.nu3
k<-7
ymin<-min(out.est[,c(1,3,7)])
ymax<-max(out.est[,c(1,3,7)])
plot(1:k,out.est[,1],xlab="Hospital de nacimiento",ylab="",ylim=c(ymin,ymax))
segments(1:k,out.est[,3],1:k,out.est[,7])

```

```
abline(h=0,col="grey70")
title("Mortalidad: efecto del hospital")

#nu5: monitoreo
out.nu5<-out.sum1b[grep("nu5",rownames(out.sum1b)),]
out.est<-out.nu5
k<-2
ymin<-min(out.est[,c(1,3,7)])
ymax<-max(out.est[,c(1,3,7)])
plot(1:k,out.est[,1],xlab="Monitoreo",ylab="",ylim=c(ymin,ymax))
segments(1:k,out.est[,3],1:k,out.est[,7])
abline(h=0,col="grey70")
title("Mortalidad: efecto del monitoreo")

#nu6: cesarea
out.nu6<-out.sum1b[grep("nu6",rownames(out.sum1b)),]
out.est<-out.nu6
k<-6
ymin<-min(out.est[,c(1,3,7)])
ymax<-max(out.est[,c(1,3,7)])
plot(1:k,out.est[,1],xlab="Causa de cesárea",ylab="",ylim=c(ymin,ymax))
segments(1:k,out.est[,3],1:k,out.est[,7])
abline(h=0,col="grey70")
title("Mortalidad: efecto de la causa de cesárea")

#nu4: cesarea*forma
out.nu4<-out.sum1b[grep("nu4",rownames(out.sum1b)),]
out.est<-out.nu4
k<-2
out.est <- data.frame(out.est[,c(1,3,7)])
out.est$aux1 <- substr(rownames(out.est), 9, 11)
out.est$forma1 <- as.numeric(substr(out.est$aux1, 1, 1))
out.est$causa<- substr(out.est$aux1, 3, 4)
causa <- c("DCP", "Electiva", "Emer. Fetal", "Emer. Mat", "Iterativa", "No hubo")

for(i in 1:6){
  out.est1 <- out.est[out.est$causa==i, ]
  ymin<-min(out.est1[,c(1:3)])
  ymax<-max(out.est1[,c(1:3)])
  plot(1:k,out.est1[,1],xlab="Forma de nacimiento",ylab="",ylim=c(ymin,ymax))
  segments(1:k,out.est1[,2],1:k,out.est1[,3])
  abline(h=0,col="grey70")
  title(paste(causa[i], sep=""))
}
```