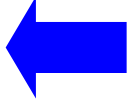


Introducción a los métodos estadísticos

bayesianos en Ecología

Pablo Inchausti

Programa del curso:

1. Introducción general
2. Elementos básicos del análisis bayesiano
3. Análisis bayesiano I
4. Análisis bayesiano II 
5. Modelos bayesianos jerárquicos

4. Análisis bayesiano II

Práctico 03

- a. Proporciones y la distr Beta.
- b. GLM de proporciones
- c. Conteos con “exceso” de ceros
- d. GLM para conteos con “exceso” de ceros

a. Proporciones y la distr Beta.

Las proporciones son valores reales en $[0,1]$ que surgen a partir de **datos binarios o conteos**, o como **registros continuos** (Douma & Weldon 2019).

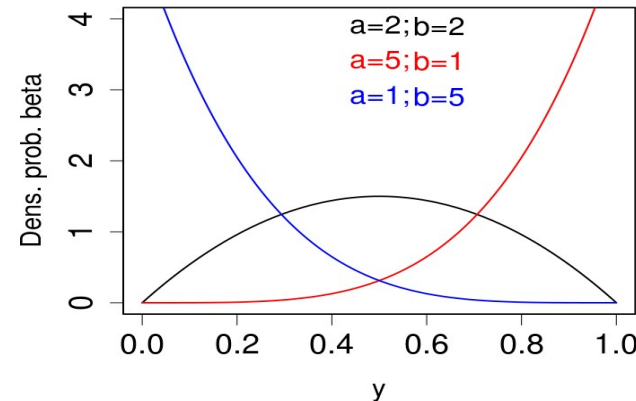
1) Suelen ser modeladas como datos binarios agrupados o como tasas (conteos/esfuerzo) con un offset.

2) Deben ser modeladas usando la distr Beta (a,b).

$$Pr(Y \leq y) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1} (1-y)^{b-1}$$

$$E(Y) = \frac{a}{a+b} = \mu_Y \quad \sigma_Y^2 = \frac{ab}{(a+b)^2(a+b+1)}$$

(la media y la varianza NO son independientes)



$Y \sim \text{Beta}(a, b)$ se re-expresa como $\text{Beta}(\mu_Y, \phi)$

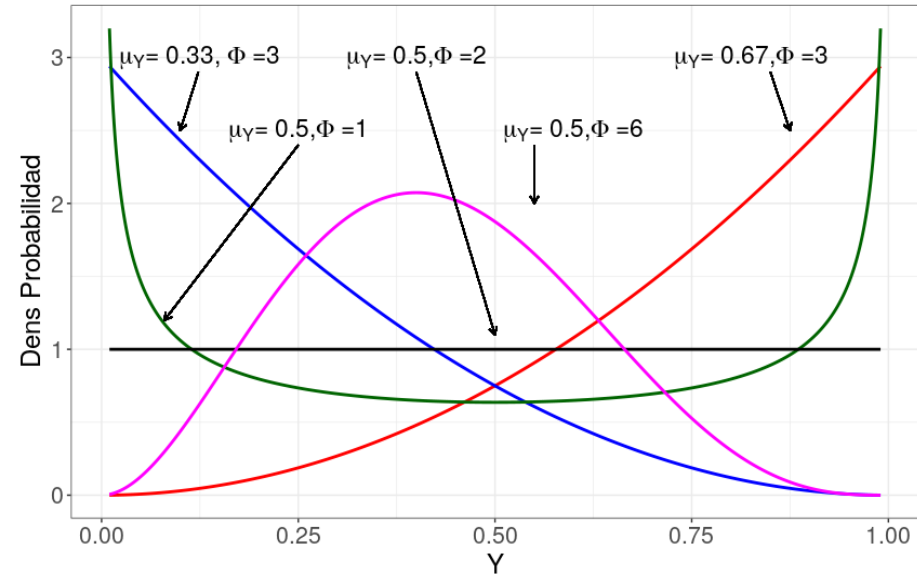
donde: $E(Y) = \mu_Y$ $\sigma_Y^2 = \frac{\mu_Y(1-\mu_Y)}{(\phi+1)}$ $\phi = a+b$ usando la función

de enlace $\text{logit}(\mu_Y) = X\beta$

$$\text{logit}(\mu_Y) \equiv \log\left(\frac{\mu_Y}{1-\mu_Y}\right)$$

Como la función de enlace no está definida en $\{0,1\}$,

se hace: $Y^* = \frac{Y(n-1) + 1/2}{n}$



La distr Beta también puede modelar var. de respuesta

reales acotadas en $[\min, \max]$: $Y^* = \frac{Y - \min(Y)}{\max(Y) - \min(Y)}$

b. GLM de proporciones

Oñatibia et al. (2008): efecto del pastoreo de ovejas sobre la cobertura de comunidades herbáceas en paisajes (# y tamaño de parches, distancia entre parches) en 3 zonas de diferente productividad primaria.

```
> DF=read.csv ("Plant cover Pr03.csv", header=T)
> summary(DF)
```

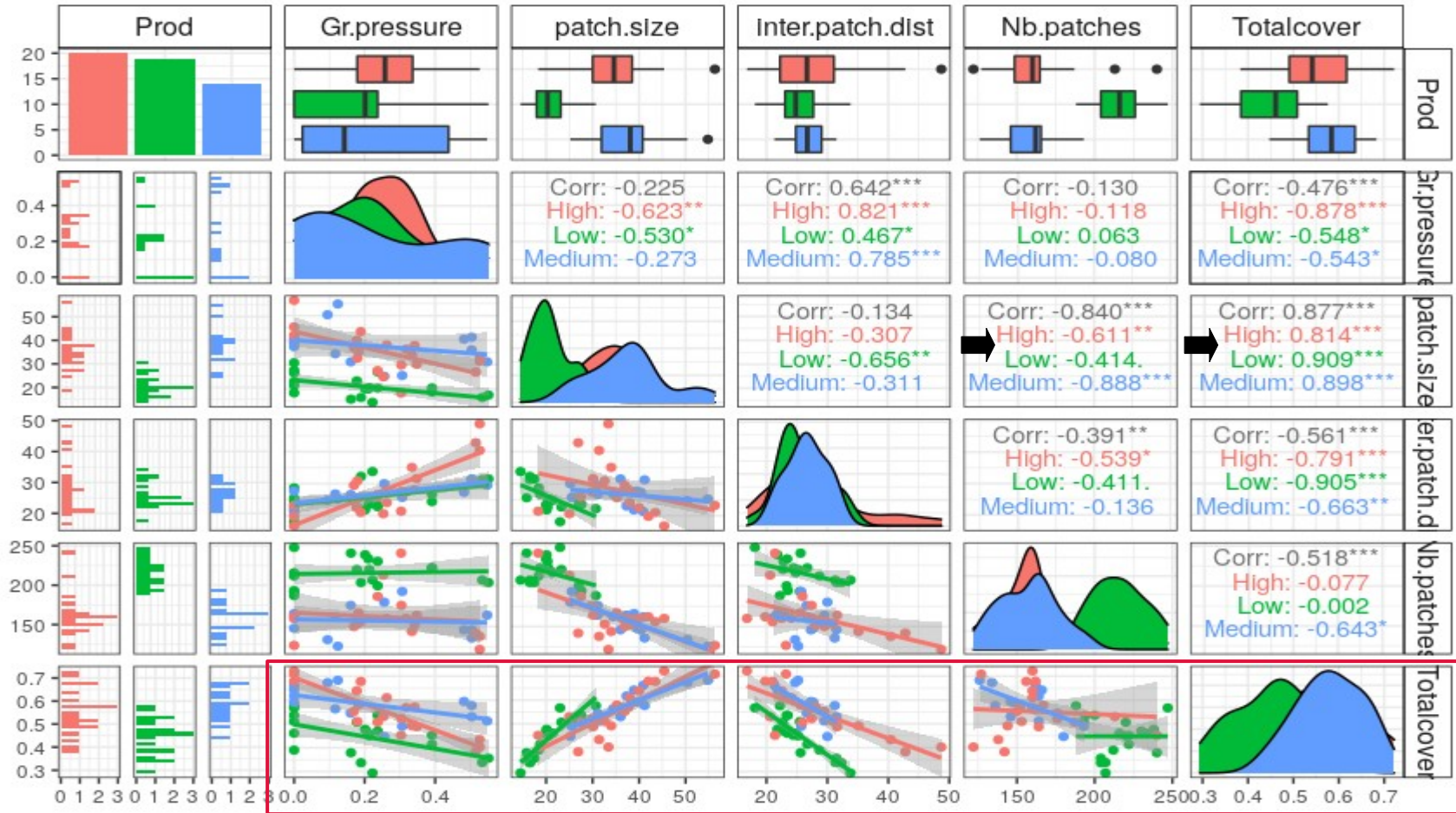
Prod	Gr.pressure	patch.size	inter.patch.dist	Nb.patches	Totalcover
High :20	Min. :0.000	Min. :14.5	Min. :16.8	Min. :122	Min. :0.294
Low :19	1st Qu.:0.095	1st Qu.:21.4	1st Qu.:23.0	1st Qu.:153	1st Qu.:0.452
Medium:14	Median :0.211	Median :30.7	Median :25.6	Median :167	Median :0.522
	Mean :0.226	Mean :30.7	Mean :26.8	Mean :180	Mean :0.522
	3rd Qu.:0.334	3rd Qu.:37.8	3rd Qu.:29.5	3rd Qu.:207	3rd Qu.:0.579
	Max. :0.551	Max. :56.3	Max. :48.7	Max. :247	Max. :0.722

```
DF.s=data.frame(scale(DF[ c("Gr.pressure","patch.size","inter.patch.dist")], scale=T, center=T),
  Totalcover=DF$Totalcover, Prod=DF$Prod)
```

← estandarización

Análisis exploratorio de datos:

Conclusions



ggpairs(data = Df, mapping = aes(col = Prod),
lower = list(continuous = wrap("smooth",
method = "lm")) + theme_bw() + etc

Modelo estadístico:

Totalcover \sim Beta(μ_Y, ϕ) con $\text{logit}(\mu_Y) = X\beta$.



Las vars explicativas a considerar:

Gr.pressure*Prod, patch.size*Prod + inter.patch.dist.

Es posible incluir vars explicativas en el parámetro ϕ empleando la función de enlace adecuada: $\log(\phi) = X\beta$.

Útil para explicitar la ecuación de un modelo complejo:

```
formula.m1=bf(Totalcover~Gr.pressure*Prod + patch.size*Prod+inter.patch.dist)+  
  lf(phi~1)
```

```
formula.m2=bf(Totalcover~Gr.pressure*Prod + patch.size*Prod+inter.patch.dist)+  
  lf(phi~Gr.pressure*Prod + patch.size*Prod+inter.patch.dist)
```

Ajustaremos ambos modelos y haremos una selección.

Veamos primero las distr previas que hay que especificar...

```
> get_prior(formula.m1, data=DF.s, family='Beta')
```

prior	class	coef	group	resp	dpar
(flat)	b				
(flat)	b		Gr.pressure		
(flat)	b	Gr.pressure:ProdLow			
(flat)	b	Gr.pressure:ProdMedium			
(flat)	b	inter.patch.dist			
(flat)	b	patch.size			
(flat)	b	ProdLow			
(flat)	b	ProdLow:patch.size			
(flat)	b	ProdMedium			
(flat)	b	ProdMedium:patch.size			
student t(3, 0, 2.5)	Intercept				
student_t(3, 0, 2.5)	Intercept				phi

¿Nivel de referencia de Prod?

¿Significado de estos parámetros?

En los GLM, las dist previas se definen en las escalas en las que el modelo es lineal: $\text{logit}(\mu_Y)$ y $\log(\phi)$.

Consideremos:

```
> prior.m1 = c(set_prior("normal(0,0.5)", class = "b"),
+               set_prior("normal(logit(0.5), (logit(0.8)-logit(0.2))/4", class = "Intercept"))
```



```
> prior.m1 = c(set_prior("normal(0,0.5)",class = "b"),  
+             set_prior("normal(logit(0.5),(logit(0.8)-logit(0.2))/4", class = "Intercept"))
```

Para los efectos de las vars. explicativas: Normal(0,0.5)→
95% de los valores estarían en [-1;1] en escala logit
(recordemos que vars. explicativas fueron estandarizadas).

Para el intercepto: $\text{logit}(0.5)=0 \rightarrow$ se espera que la
cobertura visual promedio sea de 50%, con $\text{sd}=(\text{logit}(0.8)-\text{logit}(0.2))/4=0.693$ (Wan et al 2004) ya que la cobertura
estaría en un rango [20%;80%].

Como no se especifica la distr previa para ϕ , se acepta la
dist. previa “débilmente informativa” propuesta por brms.

Examinen las dist
previas para:

```
> formula.m2=bf(Totalcover~Gr.pressure*Prod + patch.size*Prod+inter.patch.dist)+  
+             lf(phi~Gr.pressure*Prod + patch.size*Prod+inter.patch.dist)
```

Los dos modelos que vamos a ajustar:



```
m1=brm(formula=formula.m1, family=Beta(link="logit", link_phi="log"),
..... warmup=1000, data=DF.s, prior=prior.m1, chains=3, iter=3000, future=T,
..... control=list(adapt_delta=0.9999, max_treedepth=15))
m2=brm(formula=formula.m2, family=Beta(link="logit", link_phi="log"),
..... warmup=1000, data=DF.s, prior=prior.m1, chains=3, iter=3000, future=T,
..... control=list(adapt_delta=0.9999, max_treedepth=15))
```

Examinen los valores de Rhat y de ESS de cada modelo **OK ?**

Para hacer selección de modelos, hay que añadir LOO...

```
> m1=add_criterion(m1, "loo")
> m2=add_criterion(m2, "loo")
> loo_compare(m1,m2, criterion="loo")
```

	elpd_diff	se_diff
m1	0.0	0.0
m2	-4.6	6.4

m1 es más simple, pero la diferencia en loo es $<$ que el se(diferencia).

Un poco arbitrariamente, yo elegiría **m1** por su simplicidad, sin que por ello m2 sea “inadecuado”.

Interpretemos los parámetros:

```
> summary(m1)
```

```
Family: beta
```

```
Links: mu = logit; phi = log
```

```
Formula: Totalcover ~ Gr.pressure * Prod + patch.size * Prod + inter.patch.dist  
phi ~ 1
```

```
Data: DF.s (Number of observations: 53)
```

```
Draws: 3 chains, each with iter = 3000; warmup = 1000; thin = 1;  
total post-warmup draws = 6000
```

Population-Level Effects:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.10	0.02	0.07	0.13	1.00	2318	3604
phi_Intercept	7.36	0.22	6.92	7.77	1.00	3577	3539
Gr.pressure	0.06	0.03	0.01	0.11	1.00	1771	2328
ProdLow	0.09	0.04	0.02	0.17	1.00	3073	3374
ProdMedium	0.03	0.02	-0.02	0.07	1.00	2189	3085
patch.size	0.35	0.02	0.31	0.39	1.00	1864	2944
inter.patch.dist	-0.21	0.01	-0.23	-0.18	1.00	2383	3347
Gr.pressure:ProdLow	-0.07	0.03	-0.12	-0.01	1.00	1917	2804
Gr.pressure:ProdMedium	-0.07	0.03	-0.12	-0.02	1.00	1942	2856
ProdLow:patch.size	0.13	0.04	0.05	0.21	1.00	2044	3052
ProdMedium:patch.size	-0.06	0.03	-0.12	-0.01	1.00	1809	2630

Population-Level Effects:

	Estimate	Est. Error
Intercept	0.10	0.02
phi_Intercept	7.36	0.22
Gr.pressure	0.06	0.03
ProdLow	0.09	0.04
ProdMedium	0.03	0.02
patch.size	0.35	0.02
inter.patch.dist	-0.21	0.01
Gr.pressure:ProdLow	-0.07	0.03
Gr.pressure:ProdMedium	-0.07	0.03
ProdLow:patch.size	0.13	0.04
ProdMedium:patch.size	-0.06	0.03

Con las vars explicativas estandarizadas, podemos evaluar su efecto relativo para cada nivel de Gr.pressure.

Hay que obtener las ecuaciones...!

High: $0.10 + 0.06Gr + 0.35patch - 0.21inter$

Medium: $(0.10 + 0.03) + (0.06 - 0.07)Gr + (0.35 - 0.06)patch - 0.21inter$

Low: $(0.10 + 0.09) + (0.06 - 0.07)Gr + (0.35 + 0.13)patch - 0.21inter$

High: $0.10 + 0.06Gr + 0.35patch - 0.21inter$

Medium: $0.13 - 0.01Gr + 0.29patch - 0.21inter$

Low: $0.19 - 0.01Gr + 0.48patch - 0.21inter$



En todo GLM que usa función de enlace logit, se puede utilizar la “regla de 4” (Gelman et al 2014) para el efecto máximo de cada var. explicativa numérica.

↳ Esto viene del max de la derivada de la función inverso logit

High: $0.10 + 0.06\text{Gr} + 0.35\text{patch} - 0.21\text{inter}$

Medium: $0.13 - 0.01\text{Gr} + 0.29\text{patch} - 0.21\text{inter}$

Low: $0.19 - 0.01\text{Gr} + 0.48\text{patch} - 0.21\text{inter}$



Medium: para patch $0.29/4 = 0.072 \rightarrow$ un cambio de una $\text{sd}(\text{DF}\$patch)$ [1] 10.2 induce un aumento máx de 0.072 en la TotalCover promedio.

```
> post.m1=as_draws_df(m1,variable = c("^b_"), regex = T)
```

```
> head(post.m1,2)
```

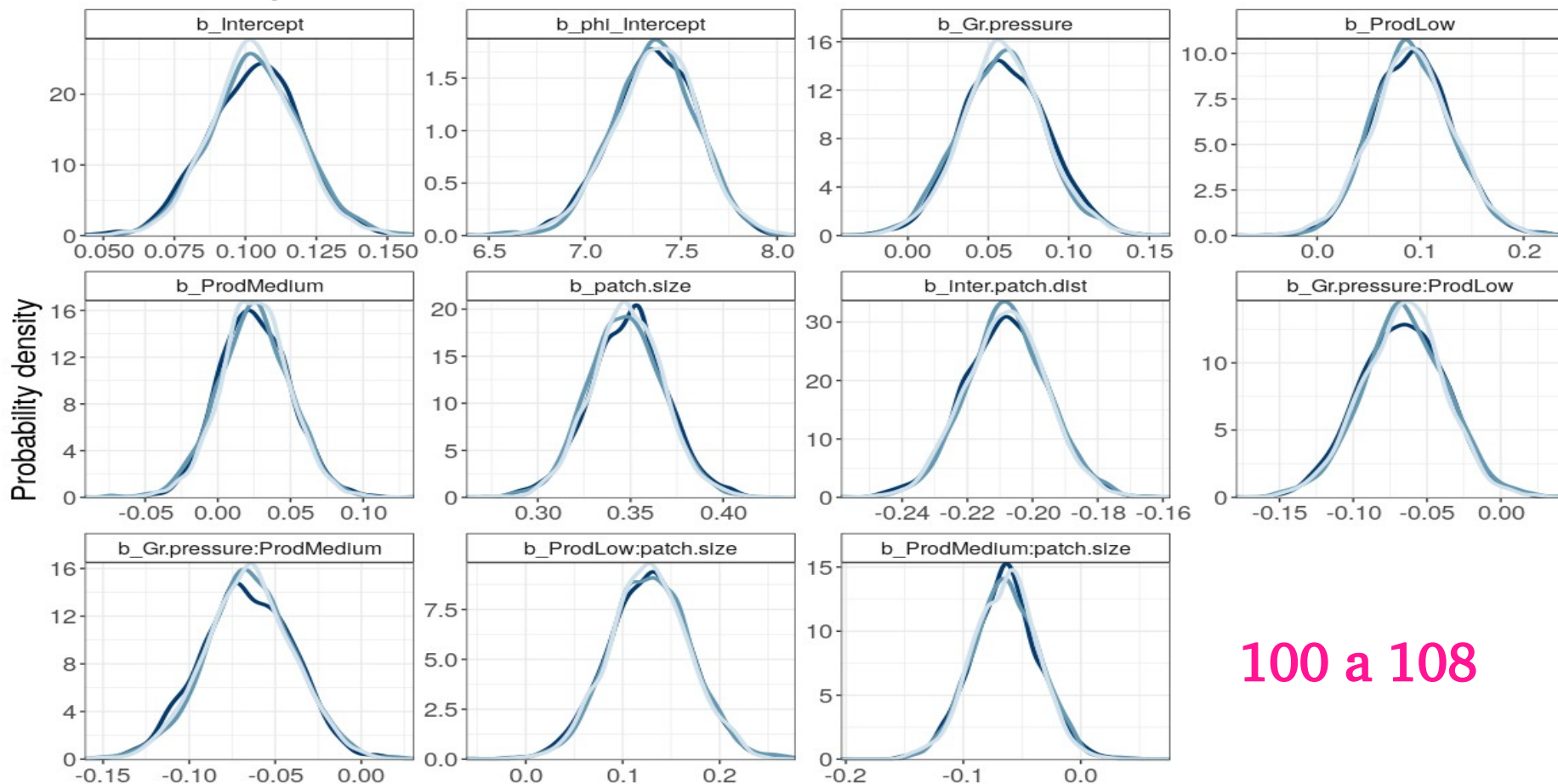
```
# A draws_df: 2 iterations, 1 chains, and 11 variables
```

	b_Intercept	b_phi_Intercept	b_Gr.pressure	b_ProdLow	b_ProdMedium	b_patch.size	b_inter.patch.dist	b_Gr.pressure:ProdLow
1	0.11	6.8	0.014	0.12	0.0537	0.35	-0.18	-0.033
2	0.10	7.5	0.052	0.13	-0.0039	0.34	-0.21	-0.036

← podemos obtener las dist posteriores.

Convergencia de las cadenas:

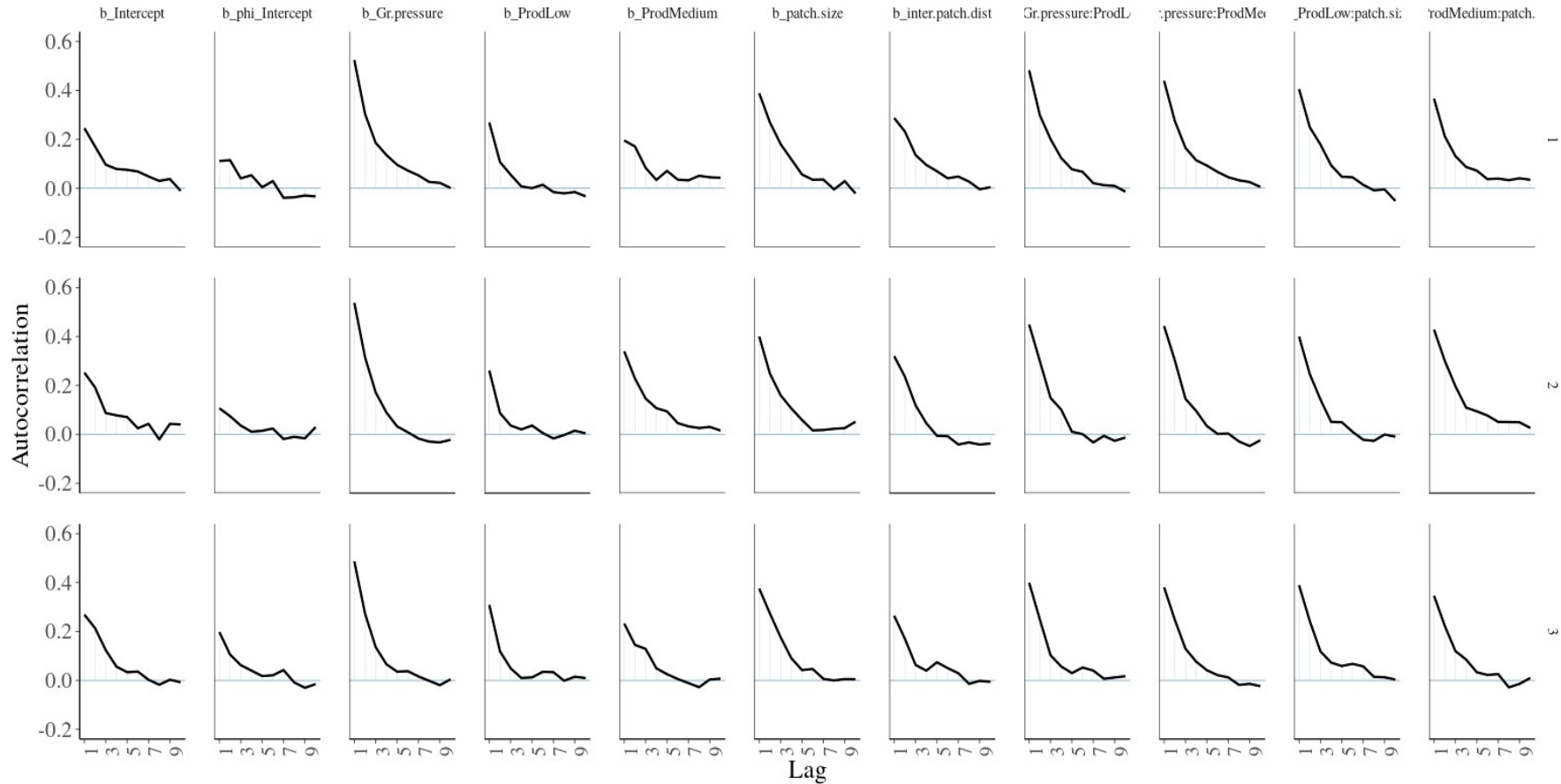
Conclusions



100 a 108

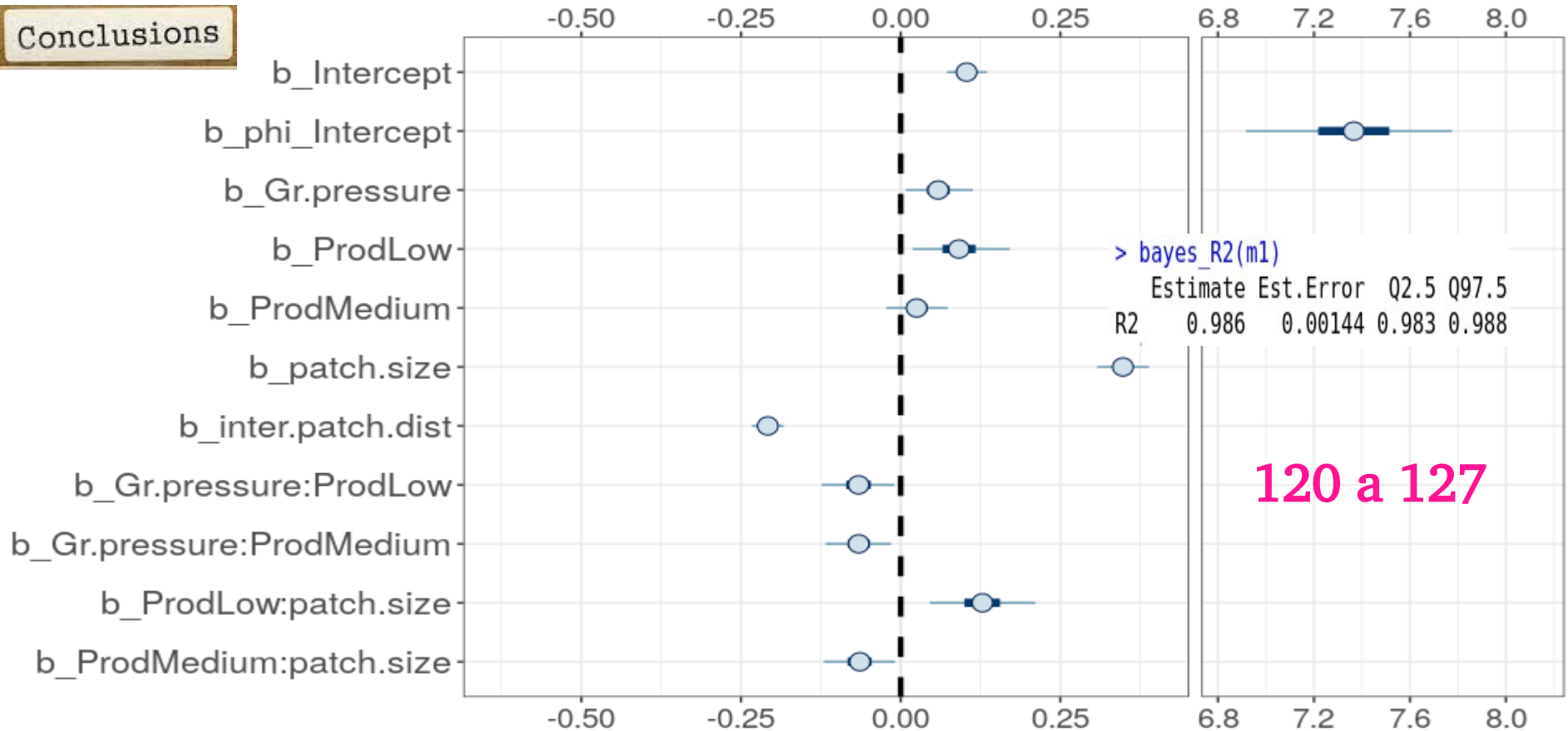
Autocorrelación de los parámetros “muestrados” por cadena:

Una solución: aumentar thin



Medias e IntCred de los parámetros del modelo:

Conclusions



Análisis de residuos:



1) Distribución predictiva posterior de m1

```
post.dist.m1=posterior_predict(m1, ndraws=1000)
```

2) Randomized quantile residuals

```
qres.m1=createDHARMA(simulatedResponse = t(post.dist.m1),  
                      observedResponse = DF.s$Totalcover,  
                      fittedPredictedResponse = apply(post.dist.m1, 2, median),  
                      integerResponse = T)
```

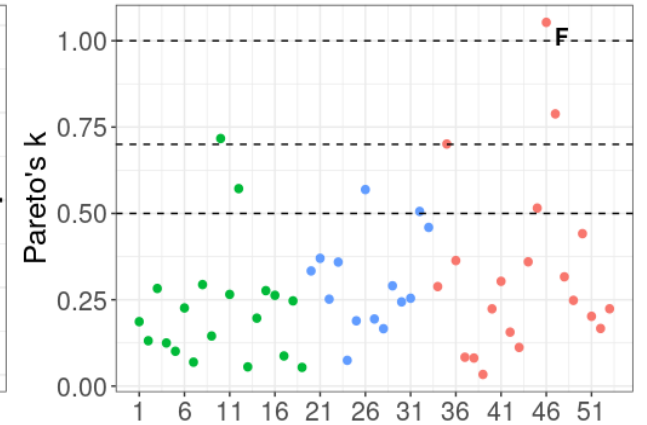
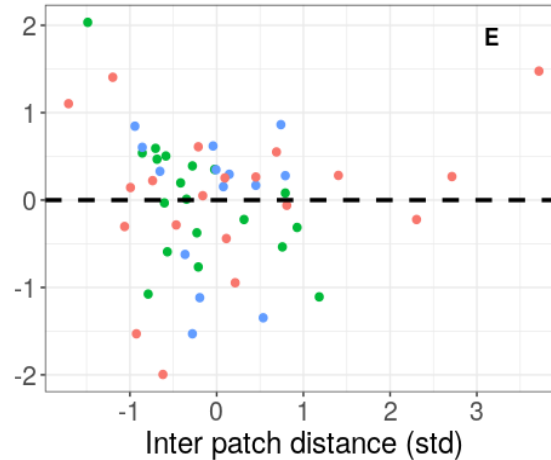
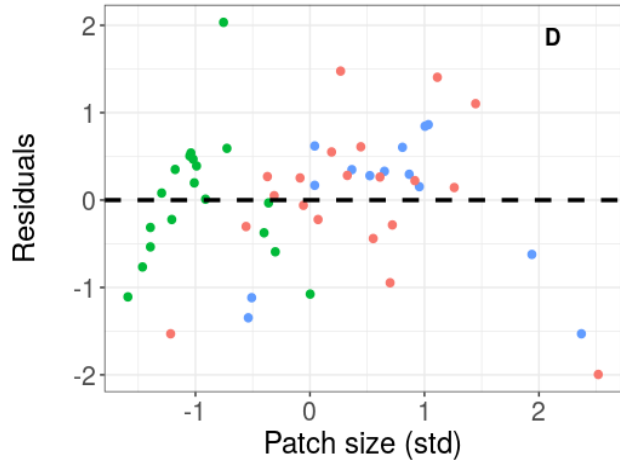
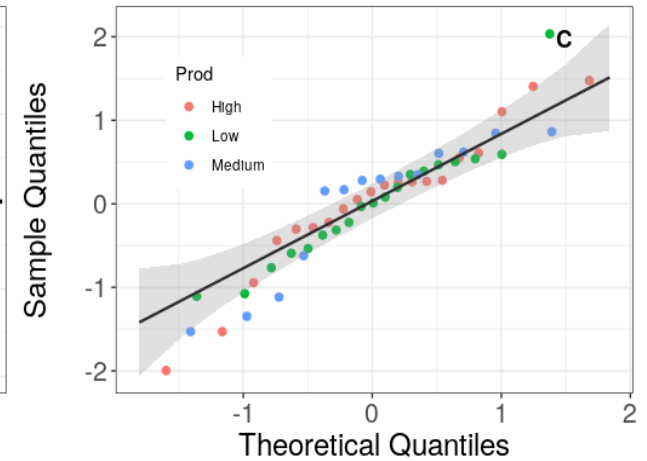
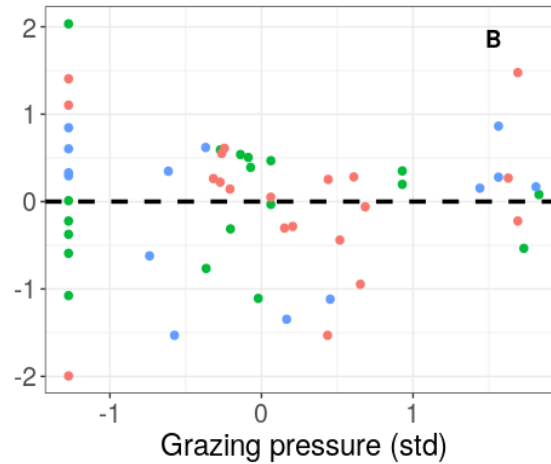
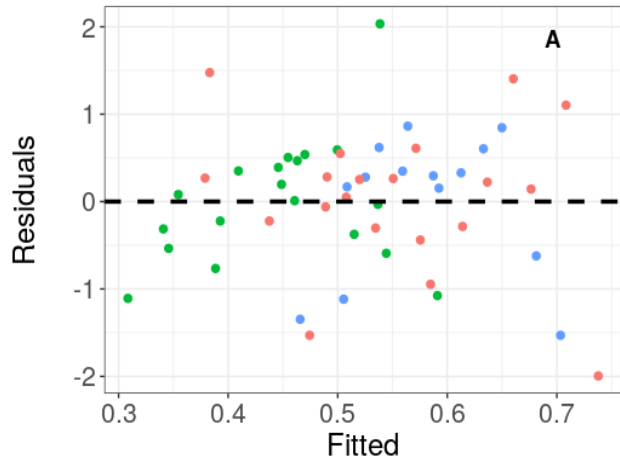
3) Convierte RQR a Normal

```
res.m1=data.frame(res=qnorm(residuals(qres.m1)))
```

4) Crea un DF con la información necesaria para los gráficos

```
res.m1=cbind(res.m1,  
             DF.s[,c("Gr.pressure", "patch.size", "inter.patch.dist", "Prod")],  
             fitted=fitted(m1, ndraws=1000)[,1], # average of fitted values  
             pareto=loo(m1, pointwise=T)$diagnostics$pareto_k)
```

Los gráficos: 150 a 210



pp-check, a partir de la distr. predictiva posterior:

```
ppc.density=ppc_dens_overlay(y=DF.s$Totalcover,yrep=post.dist.m1, trim = F,  
                             size = 0.5, alpha = 1)+
```

```
  xlab("Total Plant cover")+
```

```
  theme(axis.text = element_text(size=16),
```

```
        axis.title=element_text(size=18),
```

```
        legend.position = "none")
```

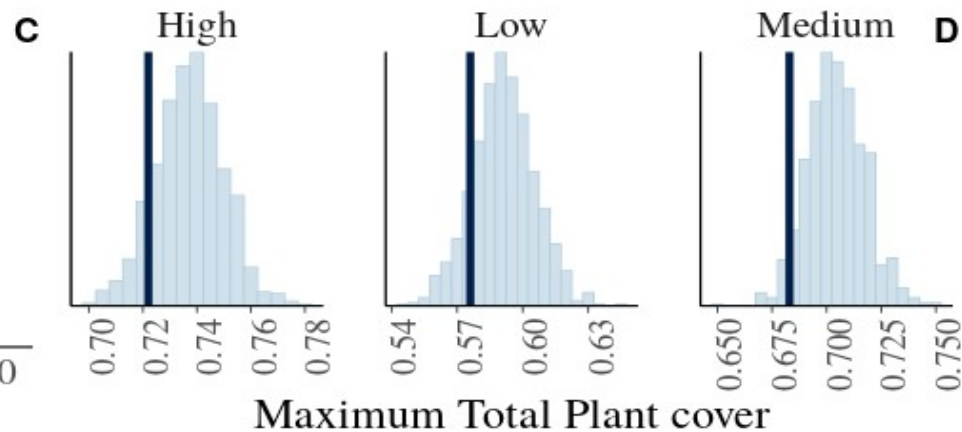
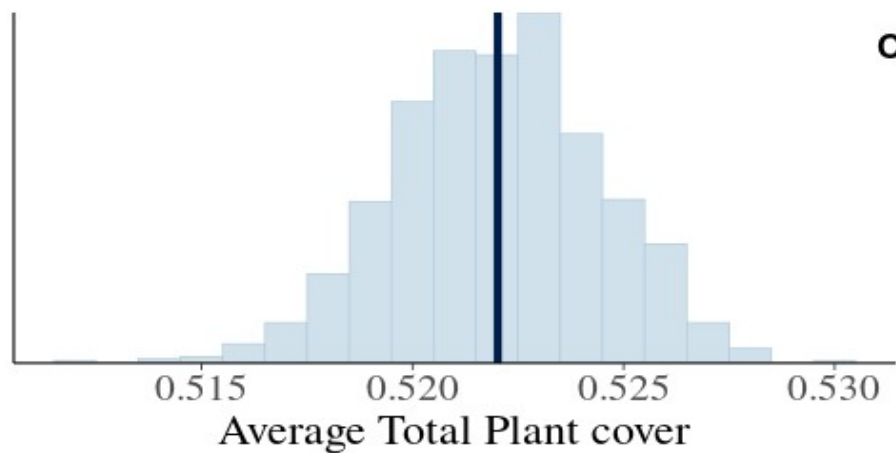
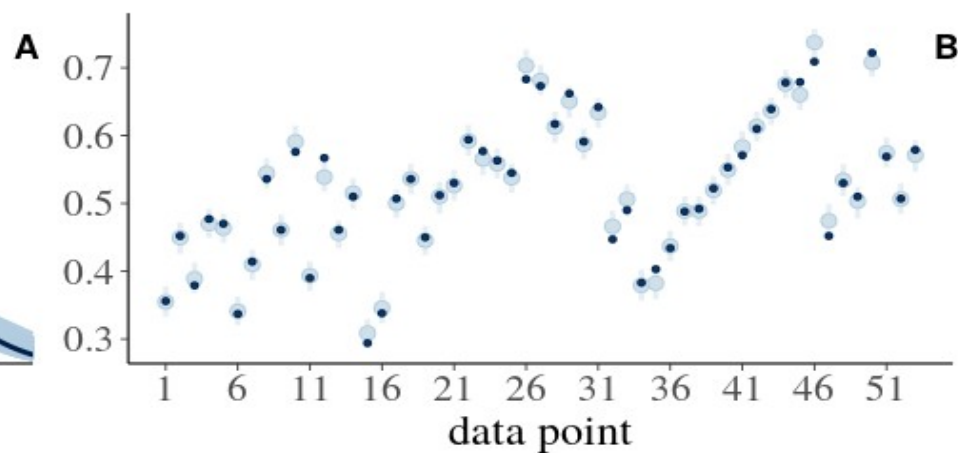
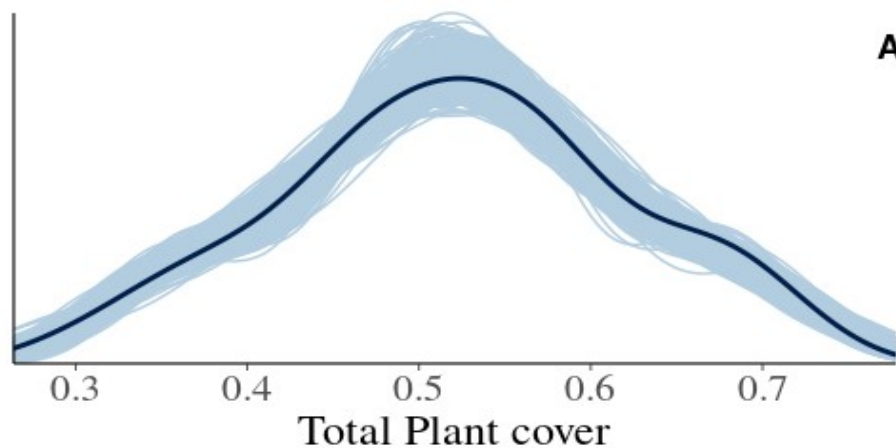
```
ppc.mean=ppc_stat(y=DF.s$Totalcover,yrep=post.dist.m1,  
                  stat = "mean", binwidth = 0.001)+ etc
```

```
ppc.interv=ppc_intervals(y=DF.s$Totalcover,yrep=post.dist.m1)+ etc
```

```
ppc_max_Prod=ppc_stat_grouped(y=DF.s$Totalcover,yrep=post.dist.m1,  
                               group = DF.s$Prod,stat = "max",binwidth=0.005)+etc
```

```
plot_grid(ppc.density, ppc.interv, ppc.mean, ppc_max_Prod, ncol=2,  
          labels = LETTERS[1:4],align="hv",label_x=0.95, label_y=0.95)
```

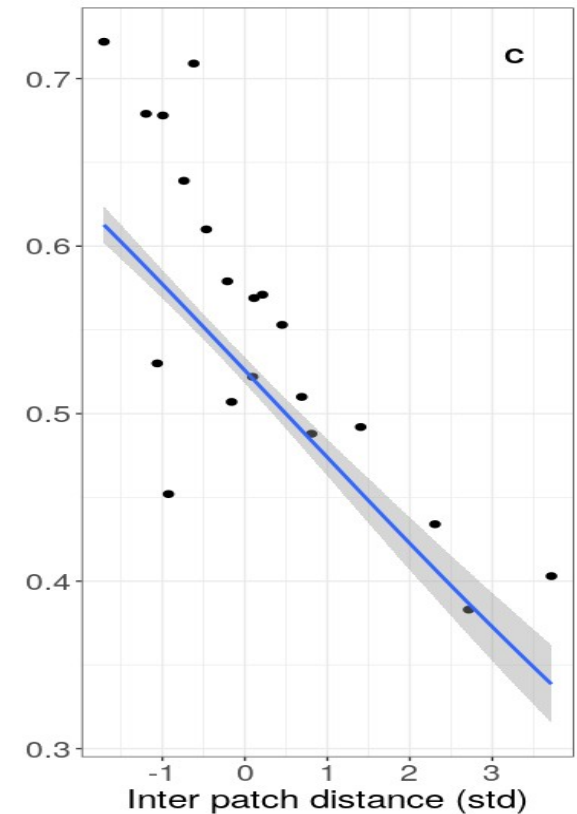
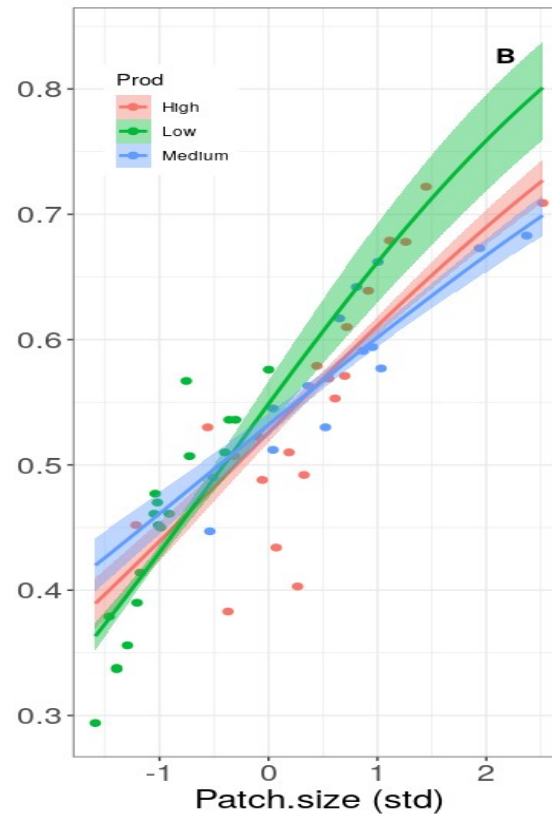
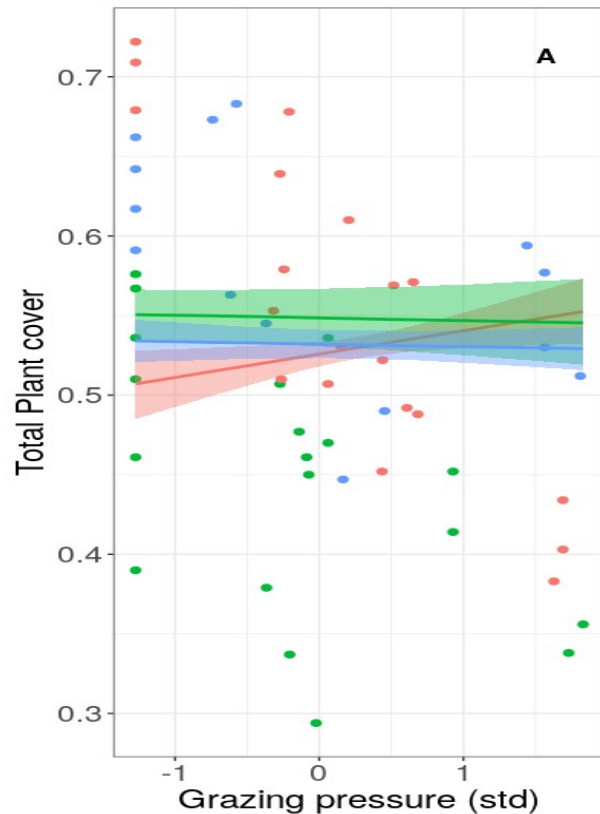




Curvas condicionales:

```
> m1.cond.eff=conditional_effects(m1)
> names(m1.cond.eff)
```

```
[1] "Gr.pressure"      "Prod"              "patch.size"
[4] "inter.patch.dist" "Gr.pressure:Prod"  "patch.size:Prod"
```



c) Conteos con exceso de ceros

Es bastante frecuente que los datos ecológicos de conteos tengan muchos ceros: ← ¿exceso de ceros?

Esto NO es una maldición o defecto de los datos, sino un aspecto de los mismos a ser modelado.



Cualquier “exceso” es respecto a una expectativa previa.

Los datos no saben que tienen ajustarse a la distr Poisson frecuentemente usada para modelar conteos.

Esta distr tiene un sólo parámetro ($\text{media}=\text{var}$) para ajustar la tendencia central y la variabilidad de los datos.

(en otra medida, lo mismo ocurre para la distr BiNeg)

Los **modelos de mezcla (mixture models)** permiten modelar directamente datos con “exceso de ceros”.

Zero Inflated: dos tipos de ceros → verdaderos (ausente) + falsos (presente pero no observada). (Lambert 1992)

Zero Augmented (hurdle): procesos diferentes dictan la presencia y dada presencia, la abundancia. (Mullahy 1986)

La **distinción entre ZI y ZA puede ser sutil** y se realiza en base al diseño y a los objetivos del estudio.

Los **modelos de mezcla** se componen de dos partes:

1. Los conteos → Poisson o BiNeg , link=log.
2. Los ceros → Binomial, link=logit.

El ZIPoisson (podría ser ZIBi) se puede visualizar como:

a. $Z \sim \text{Binomial}(1, \pi)$

b. $Y \sim \text{Poisson}(\mu * Z)$

$Y \sim ZIP(\pi, \mu)$ **Y**

$\xrightarrow{\pi} 0$ $Pr(Y_i = 0 | X_i) = \pi_i + (1 - \pi) * e^{-\mu}$ $\text{logit}(\pi) = X \beta$

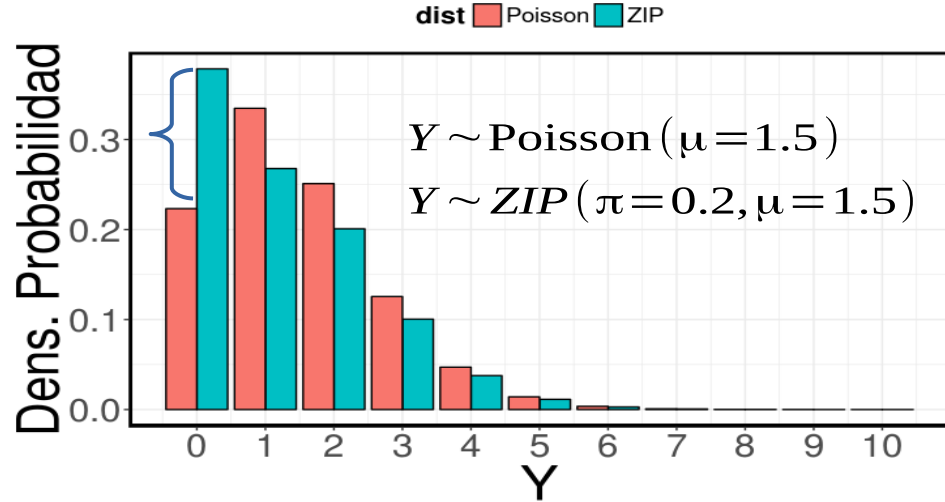
$\xrightarrow{1 - \pi} Y$ $Pr(Y_i = y_i | X_i) = (1 - \pi) \left[\frac{\mu^{y_i} e^{-\mu}}{y_i!} \right]$ $\log(\mu) = X \beta$

$E(Y_i) = (1 - \pi) \mu$ $Var(Y_i) = (1 - \pi)(\mu + \pi \mu^2)$

Las **var. explicativas** que determinan $\text{logit}(\pi)$ y $\log(\mu)$ pueden o no ser las mismas .

Esencialmente , los modelos de mezcla ZI y ZA son “dos GLM anidados” formando un solo modelo.

Comparación entre Poisson y ZI-Poisson:



ZA o Modelos condicionales: los ceros son modelados como distr. Binomial y el resto de los valores por una distr. Poisson (o Bi-) truncada.

ZA-Poisson

$$\begin{array}{c}
 \pi \nearrow \\
 \mathbf{Y} \\
 \nwarrow 1-\pi
 \end{array}$$

$$\begin{array}{c}
 \mathbf{0} \\
 \mathbf{Y|Y>0}
 \end{array}$$

$$Pr(Y=0|X)=\pi \quad \text{logit}(\pi)=X\beta$$

$$Pr(Y_i=y_i|X_i)=\frac{(1-\pi)}{(1-e^{-\mu})} \frac{\mu^{y_i} e^{-y_i}}{y!}$$

$$\log(\mu_i)=X\beta$$

d. GLM para conteos con “exceso” de ceros

Ej. efecto de la precipitación y la densidad sobre la fecundidad de un paseriforme en una isla en California (Sofaer et al 2014).



```
> DF1=read.csv("warblers Pr03.csv",header=T)
> summary(DF1)
```

Year	NumFledged	BreedingDensity	Precip
Min. :2003	Min. :0.0	Min. :3.43	Min. : 9.1
1st Qu.:2004	1st Qu.:0.0	1st Qu.:4.93	1st Qu.:18.5
Median :2005	Median :0.0	Median :4.95	Median :20.6
Mean :2006	Mean :1.4	Mean :5.23	Mean :30.4
3rd Qu.:2007	3rd Qu.:3.0	3rd Qu.:5.59	3rd Qu.:39.2
Max. :2009	Max. :6.0	Max. :6.32	Max. :62.0

```
> table(DF1$NumFledged)
```

0	1	2	3	4	5	6
95	3	28	32	19	2	2

```
> DF1$Year=as.factor(DF1$Year)
```

```
> table(DF1$NumFledged==0)/length(DF1$NumFledged)
```

FALSE	TRUE
0.475	0.525

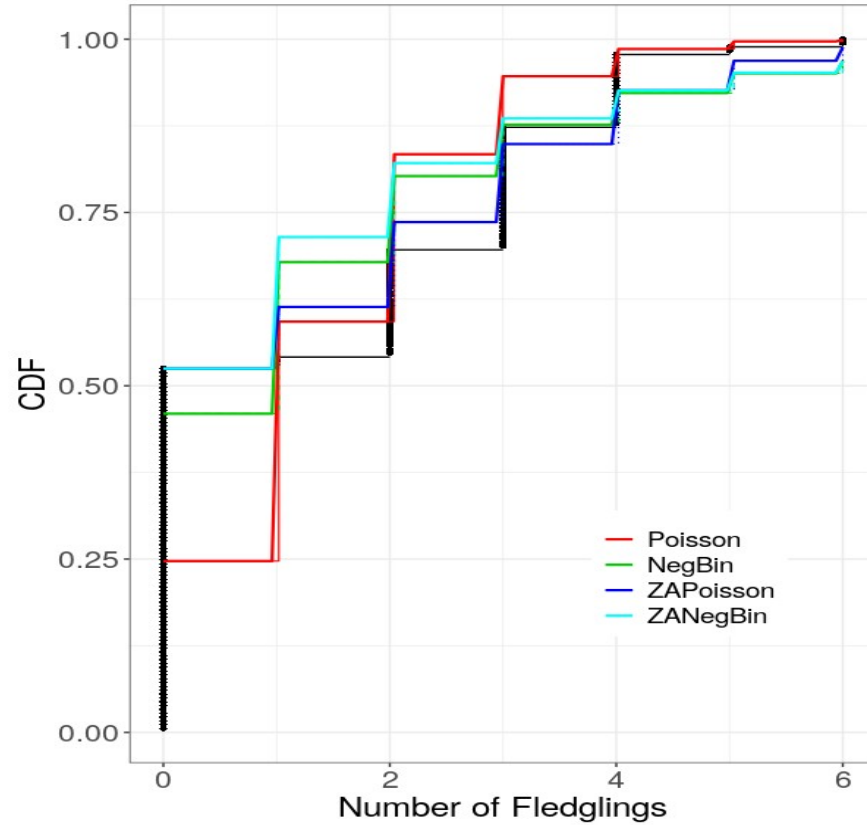
EXCESS?

**Sin ninguna justificación,
vamos ajustar modelos ZA.**

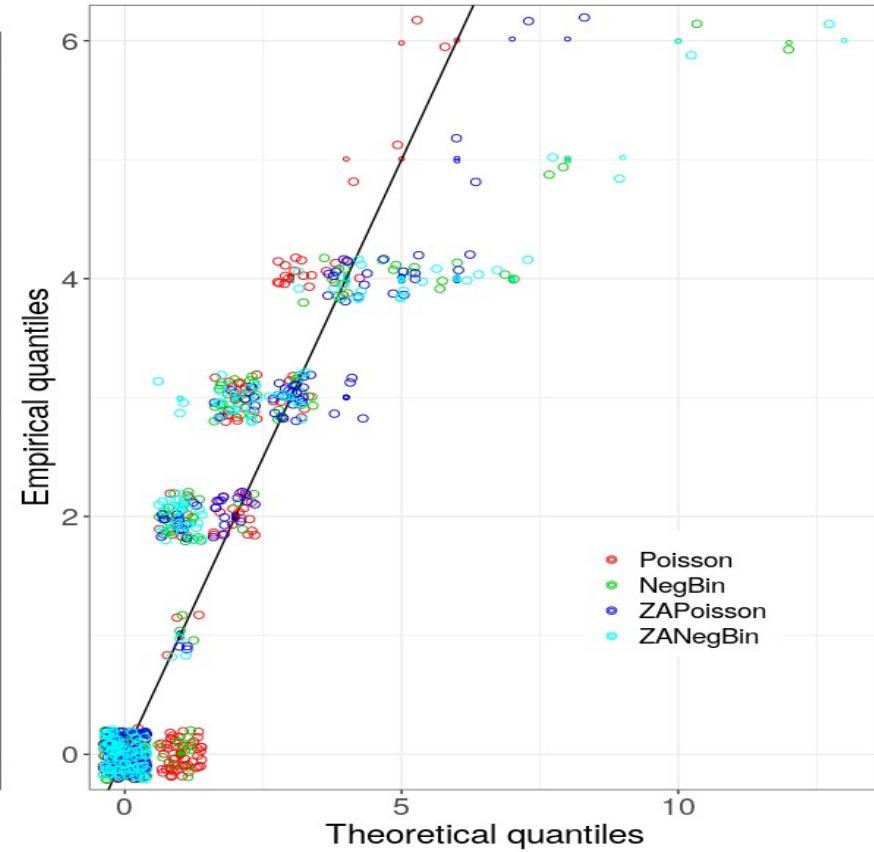
Distr. probabilidades de la var respuesta:

`poiss=fitdist(DF1$NumFledged,"pois")` $Y \sim \text{Poisson}(\mu_Y)$
`negbin=fitdist(DF1$NumFledged,"nbinom")` $Y \sim \text{Binomial negativa}(\mu_Y, \phi)$
`ZAP = fitdist(DF1$NumFledged, "ZAP",discrete=T,` $Y \sim \text{Zero Aug. Poisson}(\mu_Y, \pi)$
 `start = list(mu=mean(DF1$NumFledged),`
 `sigma=mean(DF1$NumFledged == 0)))` ←
`ZANBI = fitdist(DF1$NumFledged, "ZANBI",` $Y \sim \text{Zero Aug. Binomial Neg}(\mu_Y, \phi, \pi)$
 `start = list(mu=mean(DF1$NumFledged),`
 `sigma=(mean(DF1$NumFledged) ^ 2) /` ←
 `(var(DF1$NumFledged)-mean(DF1$NumFledged)),`
 `nu=mean(DF1$NumFledged == 0)),` ←
 `method = "mge",optim.method = "L-BFGS-B",`
 `lower = c(0.01, Inf), upper = c(0, Inf))`

Estos comandos sólo estiman los parámetros de las distr.
de probabilidades necesarios para los gráficos.



281 a 299

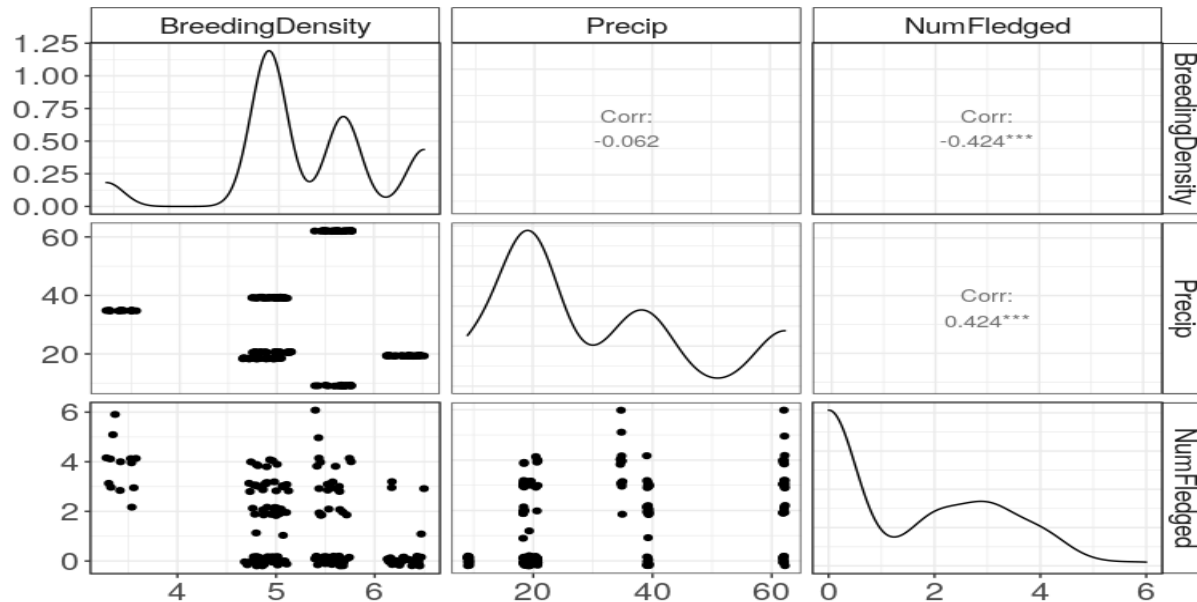


¿Diría ZAPoisson?

Modelo estadístico: $Y \sim \text{Zero Aug. Poisson}(\mu_Y, \pi)$

con $\log(\mu_Y) = X\beta$, $\text{logit}(\pi) = X\beta$

Se puede (o no) incluir vars explicativas en cada parte.



¿Qué otros gráficos exploratorios se podrían hacer?

```
ggpairs(Df1[,c("BreedingDensity", "Precip", "NumFledged")], lower = list(continuous =  
  wrap("points", position=position_jitter(height=0.2, width=0.2))) + etc
```

302-308

Usando las fórmulas para ver las distr previas a especificar:

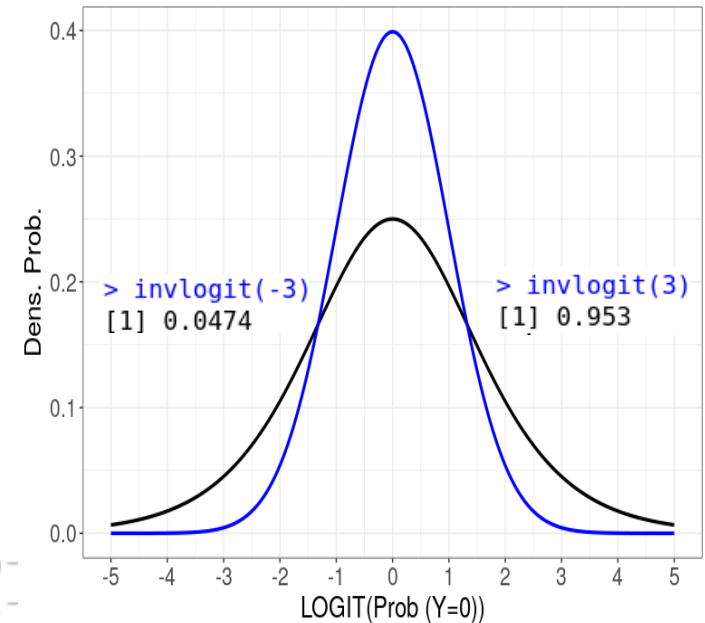
```
formula.m3=bf(NumFledged~BreedingDensity+Precip)+~  
.....lf(hu~1)-  
formula.m4=bf(NumFledged~BreedingDensity+Precip)+~  
.....lf(hu~BreedingDensity+Precip)-
```



323-330

```
> get_prior(formula=formula.m4,data=DF1s,family='hurdle_poisson')
```

	prior	class	coef	group	resp	dpar
	(flat)	b				
	(flat)	b	BreedingDensity			
	(flat)	b	Precip			
	student_t(3, -2.3, 2.5)	Intercept				
	(flat)	b				hu
	(flat)	b	BreedingDensity			hu
	(flat)	b	Precip			hu
	logistic(0, 1)	Intercept				hu



Vamos a especificar las distr previas:

```
prior.m3=c(set_prior("normal(0,0.5)",class="b",coef="Precip"))-  
prior.m4=c(set_prior("normal(0,0.5)",class="b",coef="Precip"),-  
.....set_prior("normal(0,0.45)",class="b",coef="Precip",dpar="hu"))-
```

Vamos a ajustar los siguientes dos modelos:

```
m3=brm(formula=formula.m3,data=DF1s,family='hurdle_poisson',prior=prior.m3,
·····warmup=1000,chains=3,iter=2000,thin=3)~
m4=brm(formula=formula.m4,data=DF1s,family='hurdle_poisson',prior=prior.m4,
·····warmup=1000,chains=3,iter=2000,thin=3)~
```



Y para compararlos:

```
m3=add_criterion(m3,criterio="loo")
m4=add_criterion(m4,criterio="loo")
```

```
> loo_compare(m3,m4)
      elpd_diff se_diff
m4      0.0      0.0
m3    -31.3     6.6
```

```
> summary(m4)
```

```
Family: hurdle_poisson
Links: mu = log; hu = logit
Formula: NumFledged ~ BreedingDensity + Precip
        hu ~ BreedingDensity + Precip
Data: DF1s (Number of observations: 181)
Draws: 3 chains, each with iter = 2000; warmup = 1000; thin = 3;
       total post-warmup draws = 1000
```

$$efecto_{GLM \log link} = 100 * (\exp(pend) - 1)$$

$$efecto_{GLM \logit link}^{MAX} \approx \frac{pend}{4}$$

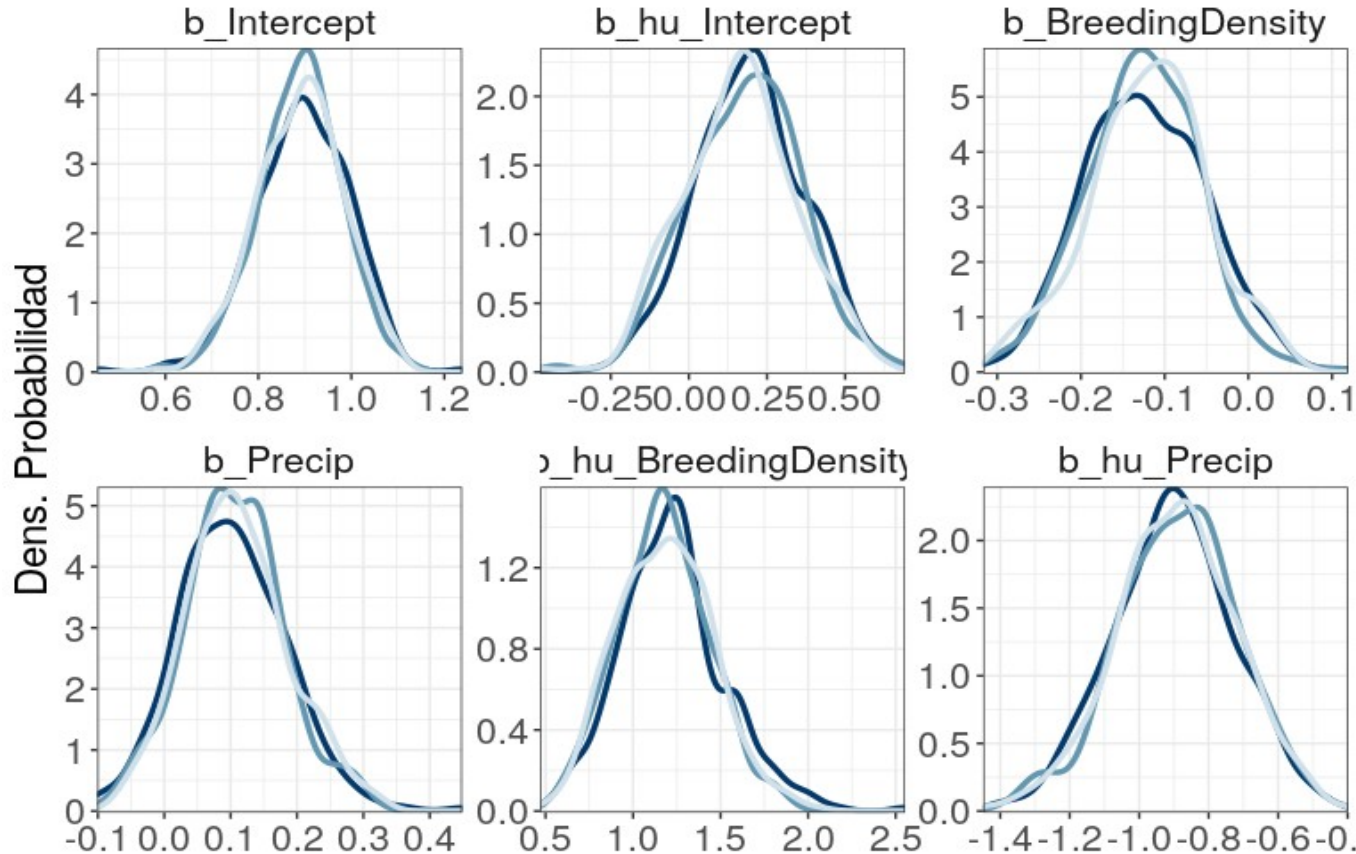
¿Interpretación?

Importancia
relativa de las
vars explic.?

Population-Level Effects:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.89	0.09	0.70	1.06	1.00	859	956
hu_Intercept	0.18	0.18	-0.15	0.51	1.00	1032	920
BreedingDensity	-0.12	0.07	-0.26	0.02	1.00	908	826
Precip	0.11	0.08	-0.04	0.27	1.00	1018	920
hu_BreedingDensity	1.19	0.28	0.69	1.79	1.00	896	947
hu_Precip	-0.89	0.17	-1.25	-0.56	1.00	920	942

Examinemos la convergencia del modelo m4:



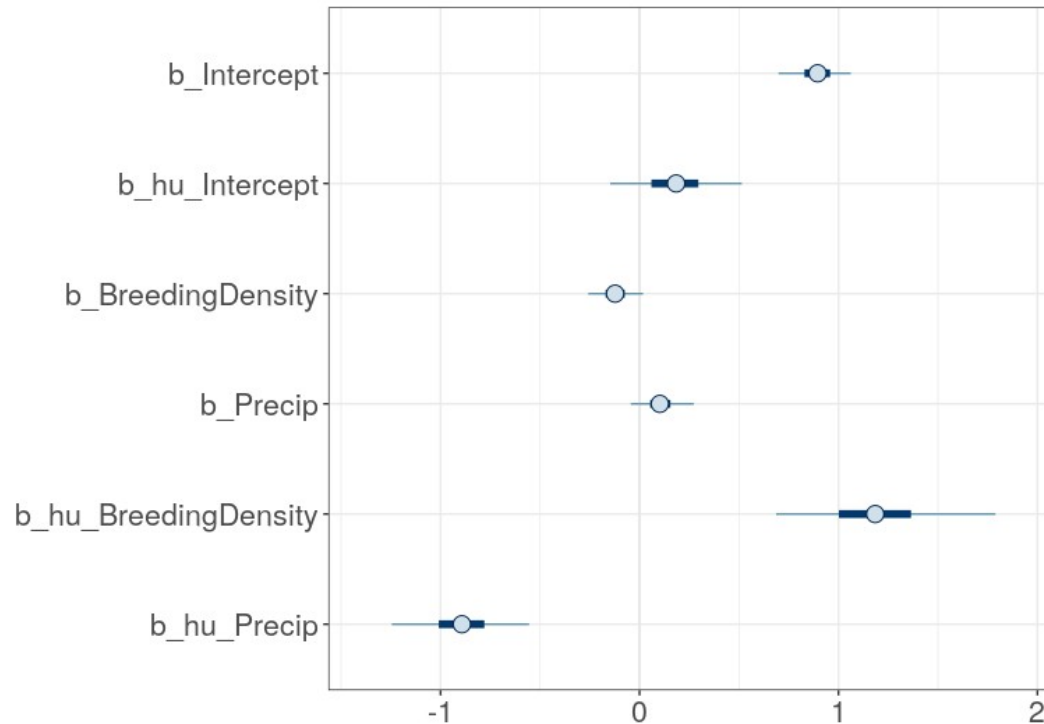
350-358



Miren los “trace plots”
y autocorrelaciones

361-379

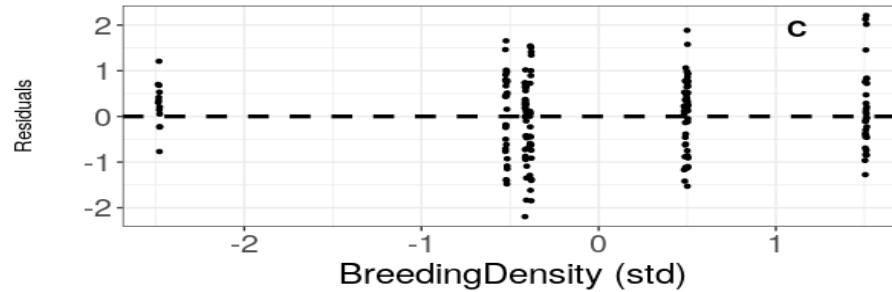
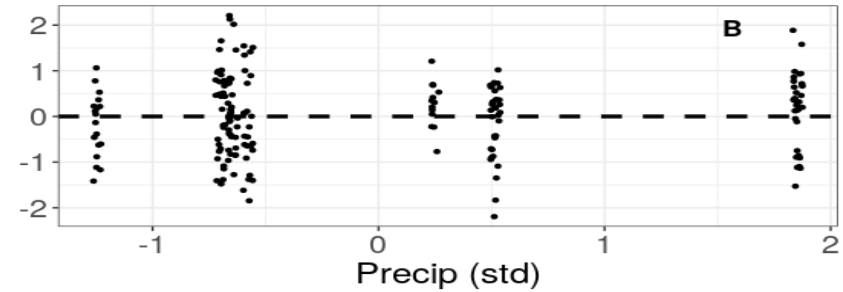
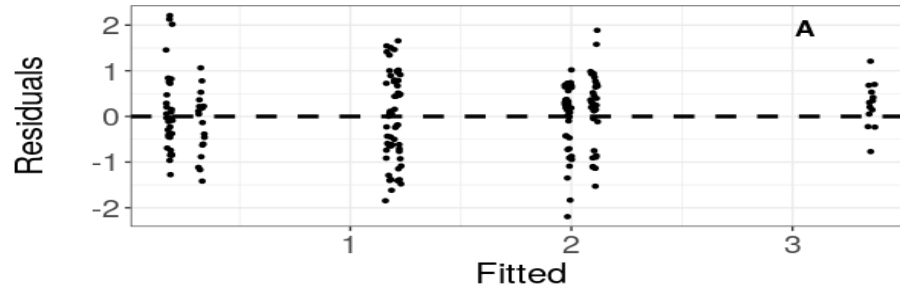
Medias e IntCred 95% de los parámetros:



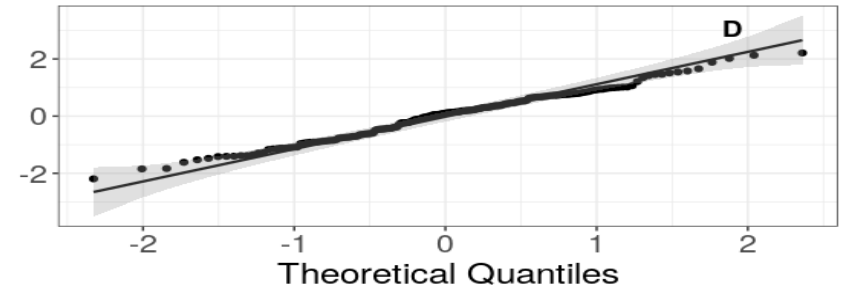
`mcmc_plot(m4, regex_pars = "^b", type="intervals", prob_outer = 0.95) +
+ etc`

Análisis de residuos (con RQR!)

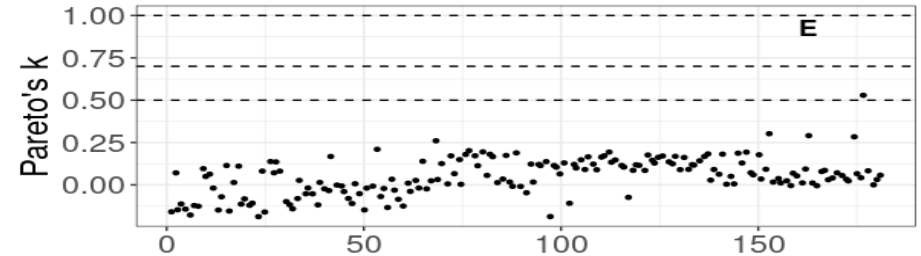
410 a 455



Sample Quantiles



1. Generar la distr pred posterior
2. Obtener los RQR
3. Convertir los RQR a dist Normal
4. Generar un DF con todo (403 a 408)



Uso de los posterior-predictive checks:

```
ppc.density.m4=ppc_dens_overlay(y=DF1s$NumFledged,yrep=dist.pred.post.m4,  
                                trim = F, size = 0.5, alpha = 1)+ etc
```

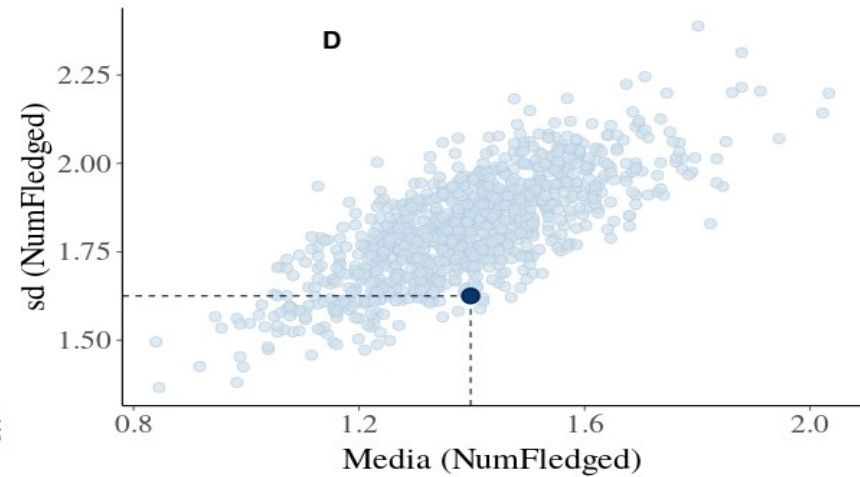
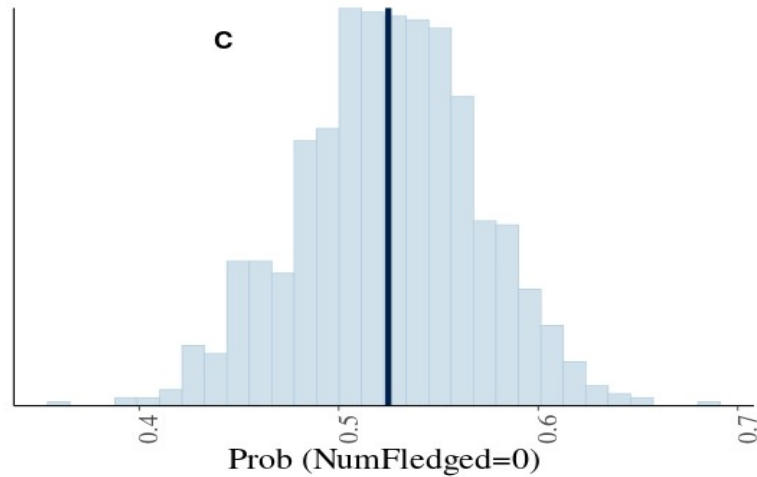
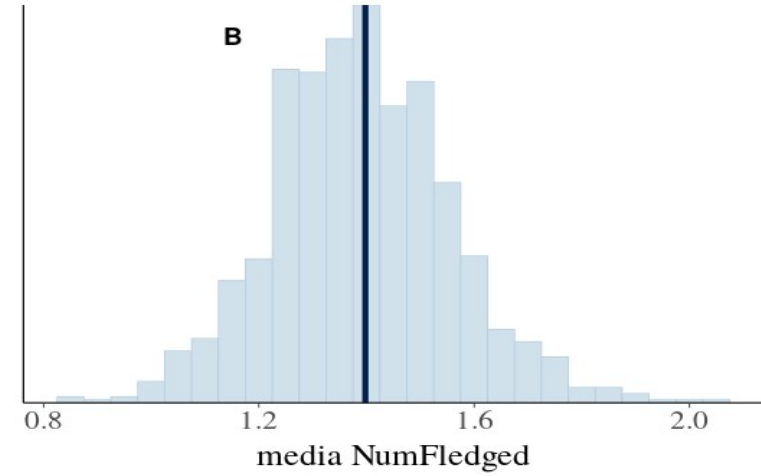
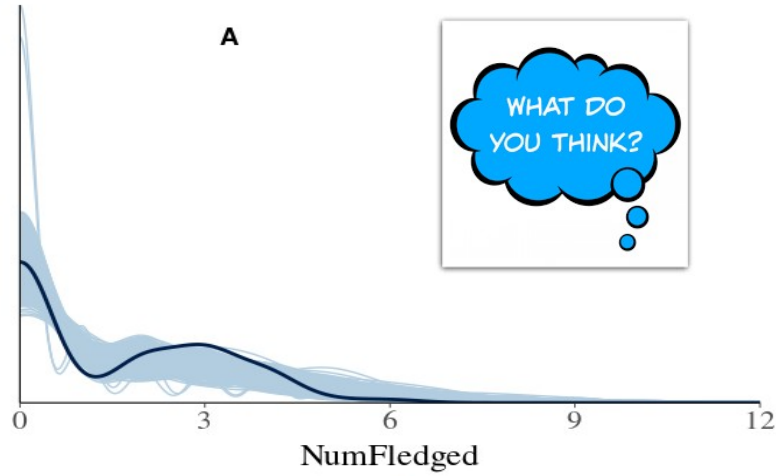
```
ppc.mean.m4=ppc_stat(y=DF1s$NumFledged,yrep=dist.pred.post.m4, stat =  
                    "mean", binwidth = 0.05)+ etc
```

```
cero=function (x){sum(x==0)/length(x)}
```

```
ppc_ceros.m4=ppc_stat(y=DF1s$NumFledged,yrep=dist.pred.post.m4,stat = cero)+  
xlab("Prob (NumFledged=0)") + etc
```

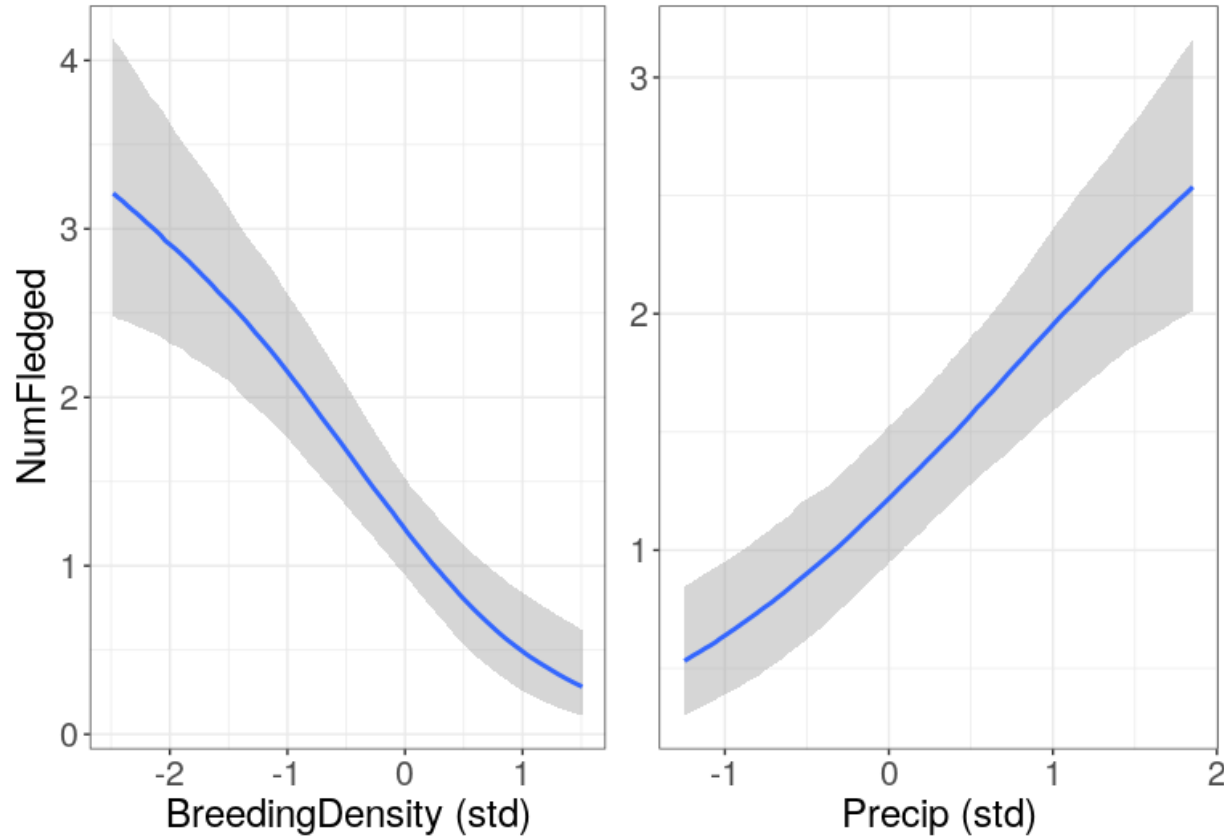
```
ppc_media_sd.m4=ppc_stat_2d(y=DF1s$NumFledged,yrep=dist.pred.post.m4,  
stat = c("mean", "sd"))+etc
```

```
plot_grid(ppc.density.m4, ppc.mean.m4,ppc_ceros.m4,ppc_media_sd.m4, ncol=2,  
          labels = LETTERS[1:4],align="hv",label_x=0.35, label_y=0.95)
```



Curvas condicionales:

```
m4.cond.eff=conditional_effects(m4)
names(m4.cond.eff) # nombre de los efectos condicionales
```



488 a 503

4. Análisis bayesiano II

Práctico 03

- a. Proporciones y la distr Beta.
- b. GLM de proporciones
- c. Conteos con “exceso” de ceros
- d. GLM para conteos con “exceso” de ceros