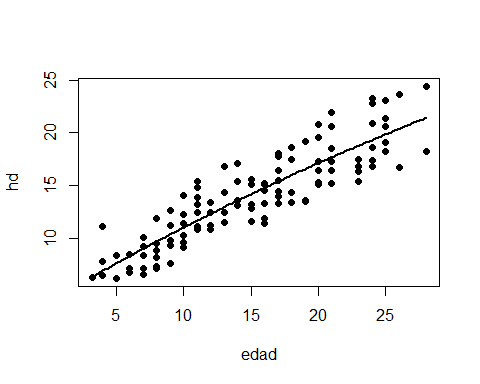
Taller5 Modelación Lineal

Juan Daniel Duque Zapata, Andrea Foronda Henao, Cristian Camilo Gañan Tapasco, Kelly Vanessa Mejía Gómez, Sofía Muñoz Duque

28 de agosto de 2019

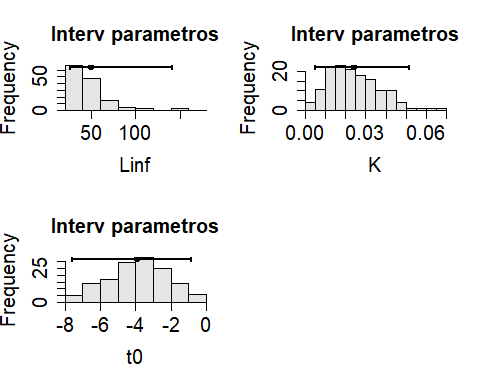
## Von Bertalanffy

library(tidyverse)  
datos<- read.csv2("taller5.csv", dec = ".")  
datos<- datos %>% select(-ï..)  
  
library(FSA)  
part<- vbStarts(hd ~ edad, data = datos)  
  
modelo<- nls(hd ~ Linf\*(1-exp(-K\*(edad-t0))), data = datos,  
 start = part)  
  
  
fitPlot(modelo)



**Figura 1: Modelo VB**

library(nlstools)  
limp<- nlsBoot(modelo, niter = 200)  
  
confint(limp,plot=TRUE,cex.lab=1.5,cex.axis=1.5,cex.main=1.5,main="Interv parametros")



**Figura 1: Modelo VB**

## 95% LCI 95% UCI  
## Linf 25.624554069 140.90734312  
## K 0.004954517 0.05182253  
## t0 -7.578263009 -0.87760341

htest(limp,"K",0.24,"less")

## Ho Value p value  
## K 0.24 0

broom::tidy(modelo)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| Linf | 57.1576873 | 49.9879912 | 1.143428 | 0.2555114 |
| K | 0.0141924 | 0.0165329 | 0.858430 | 0.3926483 |
| t0 | -5.0789771 | 2.9076422 | -1.746768 | 0.0836585 |

broom::glance(modelo)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| sigma | isConv | finTol | logLik | AIC | BIC | deviance | df.residual |
| 2.083647 | TRUE | 1e-07 | -226.7025 | 461.4051 | 472.0588 | 447.1833 | 103 |

**Tabla 1: modelo VB**

En la **figura 1**, se puede observar el ajuste del modelo, al parecer ajusta bien los datos, sin embargo, al observar las distribuciónes de los parámetros Linf, K y solo este ultimo sigue una distribución normal, el p-value para es significativo lo que quiere decir que k es una buena estimación como se dio por el modelo.

Ahora se hace preciso, mirar como se comporta el modelo si se separa por localidad, para esto, se usa la variable loc de los datos y se la hace una especie de transformación donde: y , este proceso se hizo mediante una cooerción explicita.

dato1<- datos %>% group\_by(loc) %>% nest()  
  
modelo<- function(data){  
 part<- vbStarts(hd ~ edad, data = data)  
 nls(hd ~ Linf\*(1-exp(-K\*(edad-t0))), data = data,  
 start = part)  
}  
  
dato1<- dato1 %>%   
 mutate(model= map(data, modelo))  
  
  
model2<- dato1 %>%   
 mutate(model2= map(model, broom::tidy)) %>%   
 unnest(model2, .drop = TRUE)  
  
model2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| loc | term | estimate | std.error | statistic | p.value |
| PB | Linf | 85.0043744 | 133.9095268 | 0.6347896 | 0.5278285 |
| PB | K | 0.0082183 | 0.0153303 | 0.5360818 | 0.5937604 |
| PB | t0 | -5.2978879 | 2.9333690 | -1.8060762 | 0.0756089 |
| CA | Linf | 42.8778613 | 18.7512009 | 2.2866728 | 0.0282008 |
| CA | K | 0.0261630 | 0.0183475 | 1.4259692 | 0.1624917 |
| CA | t0 | -3.1877527 | 2.3650781 | -1.3478425 | 0.1861301 |

model3<- dato1 %>%   
 mutate(model3= map(model, broom::glance)) %>%   
 unnest(model3, .drop = TRUE)  
  
model3

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| loc | sigma | isConv | finTol | logLik | AIC | BIC | deviance | df.residual |
| PB | 1.513092 | TRUE | 1.5e-06 | -121.28268 | 250.5654 | 259.3841 | 146.52469 | 64 |
| CA | 1.548274 | TRUE | 7.9e-06 | -70.82627 | 149.6525 | 156.3068 | 86.29754 | 36 |

**Tabla 2: modelos por localidad**

Haciendo modelos separados por localiadad hay un mejor ajuste, muestra de ello es la devience mas pequeña en ambos modelos por sitio que en los datos completos el mismo patron lo sigue el AIC y el BIC **Tabla 1** y **Tabla 2**

## Schumacher

El modelo presentado es de la forma , haciendo manipulaciónes matemáticas se pueden encontrar lo parametros y como sigue:

se aplica ambos lados de la ecuación.

De aqui es facil deducir, que la estimación de los parametros se puede llevar a cabo mediante una regresión lineal de la forma: :

param<- lm(log(hd) ~ I(1/edad), data= datos)   
broom::tidy(param)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| (Intercept) | 3.041847 | 0.0367947 | 82.67069 | 0 |
| I(1/edad) | -5.337759 | 0.3703406 | -14.41311 | 0 |

**Tabla 3 Parametros del modelo Schumacher**

Realmente lo importante en la **Tabla 3** es la culumna estimate pues son los parametros del modelo Schumacher(, ) iniciales.

Aqui es importante recordar que el modelo lineal es: donde pero si .

schu<- nls(hd ~ a\*exp(b\*(1/edad)), data = datos, start = list(a= exp(3.041847), b= -5.337759))   
  
broom::tidy(schu)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| a | 25.306928 | 1.000928 | 25.28347 | 0 |
| b | -7.750135 | 0.577018 | -13.43136 | 0 |

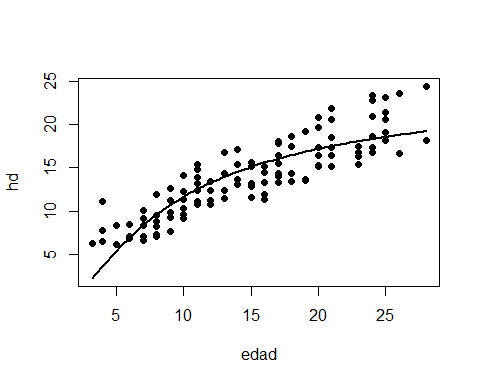
broom::glance(schu)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| sigma | isConv | finTol | logLik | AIC | BIC | deviance | df.residual |
| 2.36053 | TRUE | 1.8e-06 | -240.4399 | 486.8797 | 494.8701 | 579.4986 | 104 |

**Tabla 4: Modelo Schumacher**

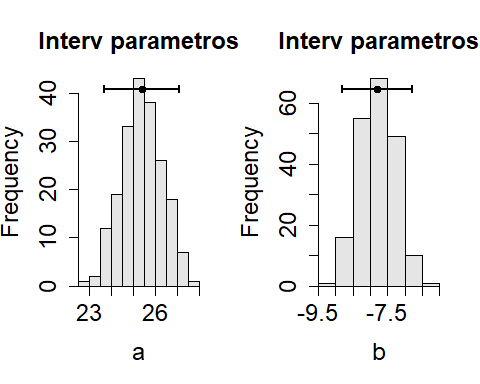
Los parametros son significativos en el modelo, ademas, según las **figuras 2 y 3** es un buen ajuste del modelo las distribuciónes de y son tienden a ser normales, tambien los errores no parecen ser heterocedásticos, sin embargo, en la **Figura 4** se observan posibles datos remotos, pero para este caso esos, podrían representar un sitio específico con caracteristicas particulares.

fitPlot(schu)



**Figura 2: Modelo Schumacher**

limp<- nlsBoot(schu, niter = 200)  
  
confint(limp,plot=TRUE,cex.lab=1.5,cex.axis=1.5,cex.main=1.5,main="Interv parametros")



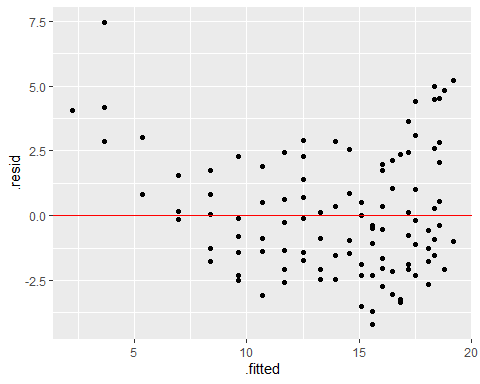
**Figura 3: Distribución de los parametros del modelo**

## 95% LCI 95% UCI  
## a 23.684417 27.086043  
## b -8.806193 -6.778085

htest(limp,"a",0.24,"less")

## Ho Value p value  
## a 0.24 1

broom::augment(schu) %>% ggplot(data = ., mapping = aes(x= .fitted, y= .resid)) + geom\_point() + geom\_abline(slope = 0, color= "red")



**Figura 4: Residuales del modelo Schumacher**

Si se comparan los modelos de Schumacher y Von Bertalanffy es posible deducir a través de sus deviance que para esto mismo pasa para y este mismo comportamiento con AIC y BIC lo que suguiere que para los datos el mejor ajuste lo brinda Von Bertalanffy (Tabla 1 y Tabla 4).

Se quiere ahora, mirar como se comporta el modelo separado por localidad:

dato1<- datos %>% group\_by(loc) %>% nest()  
  
modelo<- function(data){  
   
 param<- lm(log(hd) ~ I(1/edad), data= data)   
 nls(hd ~ a\*exp(b\*(1/edad)), data = data, start = list(a= exp(coefficients(param)[1]), b= coefficients(param)[2]))  
}  
  
dato1<- dato1 %>%   
 mutate(model= map(data, modelo))  
  
  
model2<- dato1 %>%   
 mutate(model2= map(model, broom::tidy)) %>%   
 unnest(model2, .drop = TRUE)  
  
model2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| loc | term | estimate | std.error | statistic | p.value |
| PB | a | 24.449109 | 1.0817853 | 22.60070 | 0 |
| PB | b | -8.379744 | 0.6793161 | -12.33556 | 0 |
| CA | a | 28.955600 | 1.3509787 | 21.43305 | 0 |
| CA | b | -7.889187 | 0.6444941 | -12.24090 | 0 |

model3<- dato1 %>%   
 mutate(model3= map(model, broom::glance)) %>%   
 unnest(model3, .drop = TRUE)  
  
model3

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| loc | sigma | isConv | finTol | logLik | AIC | BIC | deviance | df.residual |
| PB | 1.890621 | TRUE | 2.2e-06 | -136.72631 | 279.4526 | 286.0667 | 232.3391 | 65 |
| CA | 1.905901 | TRUE | 2.8e-06 | -79.46528 | 164.9306 | 169.9213 | 134.4009 | 37 |

**Tabla 5: Modelo por localidad Schumacher**

El modelo por localidad ajusta mejor, que sin separarlo, pues al comparar devience, AIC y BIC de los modelos por localidad y sin separar es mas alto el resultado en este ultimo. Tambien si se hace comparaciónes entre los modelos por localidad de Schumacher y Von Bertalanffy se nota el mismo patron observado en los modelos sin separación, , es decir, el modelo Von Bertalanffy es el que mejor se ajusta tambien en modelos por localidad. (**Tabla 2** y **Tabla 5**)

## Chapman-Richards

El modelo de la forma se estimó con la ayuda de la literatura, pues es un poco complicado hallar los parametros, en varios textos la modelación de altura y edad aparecen valores del orden por lo cual se intentó a apartir de esto hallar parametros para el modelo.

pr<- lm(log(hd) ~ edad, data = datos)  
  
broom::tidy(pr)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| (Intercept) | 1.8926332 | 0.0403195 | 46.94085 | 0 |
| edad | 0.0461503 | 0.0024607 | 18.75503 | 0 |

exp(log(6.3)/(1.89263315+(0.04615031\*3.2)))

## [1] 2.464752

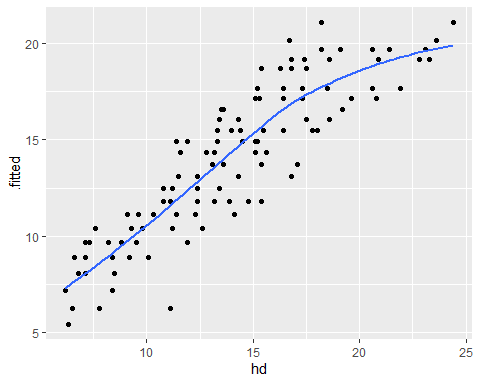
library(nlme)  
library(nls2)  
library(minpack.lm)  
library(robustbase)  
library(minpack.lm)  
  
mol<- nlsLM(hd ~ a\*(1-exp(-d\*edad))^b, data = datos, start = list(a = 30, d = 0.002, b = 0.064))  
  
broom::tidy(mol)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| a | 134.1283867 | 1357.0868153 | 0.0988355 | 0.9214610 |
| d | 0.0019729 | 0.0338334 | 0.0583118 | 0.9536132 |
| b | 0.6327532 | 0.1628066 | 3.8865335 | 0.0001801 |

broom::glance(mol)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| sigma | isConv | finTol | logLik | AIC | BIC | deviance | df.residual |
| 2.09611 | FALSE | 0 | -227.3347 | 462.6693 | 473.3231 | 452.5486 | 103 |

broom::augment(mol) %>% ggplot(data = ., mapping = aes(x= hd, y= .fitted)) +  
 geom\_point() + geom\_smooth(se= FALSE)



**Figura 5: Modelo Chapman-Richards**

Si se compara, los modelos que se hicierón anteriormente,

dato1<- datos %>% group\_by(loc) %>% nest()  
  
modelo<- function(data){  
 nlsLM(hd ~ a\*(1-exp(-d\*edad))^b, data = data, start = list(a = 30, d = 0.002, b = 0.064))  
}  
  
dato1<- dato1 %>%   
 mutate(model= map(data, modelo))  
  
  
  
model2<- dato1 %>%   
 mutate(model2= map(model, broom::tidy)) %>%   
 unnest(model2, .drop = TRUE)  
  
model2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| loc | term | estimate | std.error | statistic | p.value |
| PB | a | 159.1608375 | 1990.8979412 | 0.0799442 | 0.9365311 |
| PB | d | 0.0016097 | 0.0321282 | 0.0501027 | 0.9601966 |
| PB | b | 0.6654172 | 0.1651231 | 4.0298249 | 0.0001508 |
| CA | a | 121.1714422 | 867.3940216 | 0.1396960 | 0.8896789 |
| CA | d | 0.0030136 | 0.0369446 | 0.0815715 | 0.9354393 |
| CA | b | 0.6442113 | 0.1796630 | 3.5856658 | 0.0009901 |

model3<- dato1 %>%   
 mutate(model3= map(model, broom::glance)) %>%   
 unnest(model3, .drop = TRUE)  
  
model3

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| loc | sigma | isConv | finTol | logLik | AIC | BIC | deviance | df.residual |
| PB | 1.536393 | FALSE | 0 | -122.30659 | 252.6132 | 261.4320 | 151.07230 | 64 |
| CA | 1.566770 | FALSE | 0 | -71.28941 | 150.5788 | 157.2331 | 88.37168 | 36 |

## 

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MODELO** | | **AIC** | **BIC** | **Deviance** |
| **Von Bertalanffy** | **PB** | 250.5654 | 259.3841 | 146.52469 |
| **CA** | 149.6525 | 156.3068 | 86.29754 |
| **Schumacher** | **PB** | 279.4526 | 286.0667 | 232.3391 |
| **CA** | 164.9306 | 169.9213 | 134.4009 |
| **Chapman- Richards** | **PB** | 252.6132 | 261.4320 | 151.07230 |
| **CA** | 150.5788 | 157.2331 | 88.37168 |

## Al realizar la comparación entre los modelos propuestos por localidades, a partir de los valores obtenidos para el AIC, el BIC y la deviance, se concluye que el modelo que representa un mejor ajuste es el modelo de Von Bertalanffy, puesto que para las localidades es el que registra los menores valores de estos 3 criterios. Luego, al comparar entre el modelo de Schumacher y el modelo de Chapman-Richards se observa que este último arroja menores valores para los 3 criterios mencionados. Por esto se concluye que el modelo de Schumacher fue el modelo que presentó el menor ajuste para los modelos por localidades.

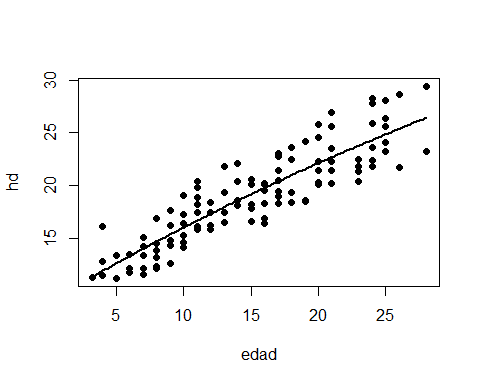
## Indices de sitio

A partir de lo concluido anteriormente, se realiza la estimación de los índices de sitio utilizando el modelo de Von Bertalanffy.

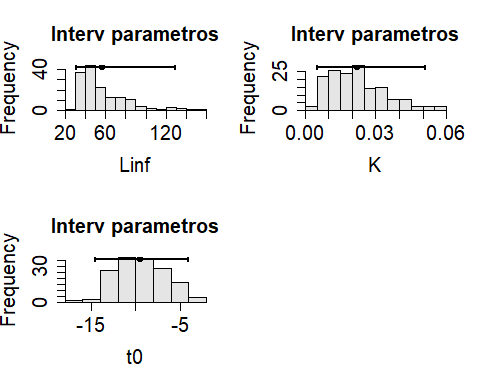
datos3<- datos %>% mutate(hd= hd+ 5)  
  
part<- vbStarts(hd ~ edad, data = datos3)  
  
modelo<- nls(hd ~ Linf\*(1-exp(-K\*(edad-t0))), data = datos3,  
 start = part)  
  
broom::tidy(modelo)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| Linf | 62.1575282 | 49.9876124 | 1.2434586 | 0.2165217 |
| K | 0.0141924 | 0.0165329 | 0.8584333 | 0.3926465 |
| t0 | -10.9878168 | 4.8230495 | -2.2781887 | 0.0247800 |

fitPlot(modelo)



library(nlstools)  
limp<- nlsBoot(modelo, niter = 200)  
  
confint(limp,plot=TRUE,cex.lab=1.5,cex.axis=1.5,cex.main=1.5,main="Interv parametros")



## 95% LCI 95% UCI  
## Linf 30.969408884 129.61738793  
## K 0.005298899 0.05122927  
## t0 -14.564125366 -4.03355695

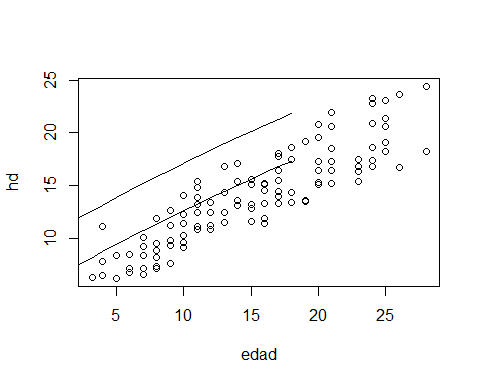
htest(limp,"K",0.24,"less")

## Ho Value p value  
## K 0.24 0

##idea  
numr<- seq(3,20,1)  
  
s<- c()  
s1<- c()  
for (i in 1:18){  
 s[i]<- 57.15\*(1-exp(-0.014\*(numr[i]+5.80)))   
}  
  
s

## [1] 6.624438 7.326867 8.019531 8.702565 9.376104 10.040278 10.695219  
## [8] 11.341054 11.977911 12.605914 13.225186 13.835849 14.438022 15.031824  
## [15] 15.617370 16.194775 16.764153 17.325616

for (i in 1:18){  
 s1[i]<- 62.16\*(1-exp(-0.014\*(numr[i]+10.98)))  
}  
  
with(datos, plot(x= edad, y= hd))  
lines(s)  
lines(s1)



Se intentó proponer una familia de curvas de índice de sitio; sin embargo estas no se obtuvieron de forma exitosa. A pesar de esto, se expone el resultado gráfico de esta tentativa, en tanto que fue complicado definir cuáles eran los criterios para poder realizar la diferenciación de los índices de sitio y la edad base para poder realizar la clasificación. Por otro lado, resultó un poco confuso el procedimiento de ajustar curvas que fueran similares al comportamiento de los datos para poder representar los difernetes sitios.