

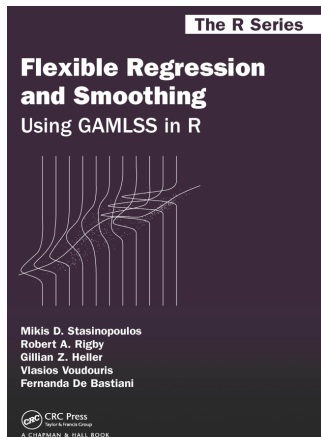
Regresión GAMLSS

Juan C. Correa

Material de uso exclusivo para
INGENIO PANTALEON, S.A.
Diagonal 6, 10-31, Zona 10

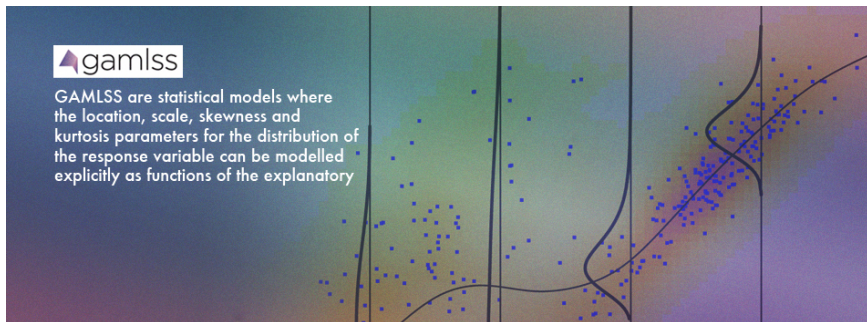
Ciudad de Guatemala

- 1 Regresión GAMLSS
- 2 Ejemplo: El Efecto Dunning-Kruger



GAMLSS es el acrónimo para Generalized Additive Models for Location, Scale, and Shape.

Regresión GAMLSS



Los cuatro parámetros que describen de forma **exhaustiva** una variable son: **Promedio, Varianza, Asimetría, y Kurtosis.**

<https://www.gamlss.com/wp-content/uploads/2018/01/DistributionsForModellingLocationScaleandShape.pdf>

Regresión GAMLSS

TABLE 6.1: Continuous distributions implemented within the **gamlss.dist** package, with default link functions

Distribution	gamlss name	Range R_Y	Parameter link functions			
			μ	σ	ν	τ
beta	BE	$(0, 1)$	logit	logit	-	-
Box-Cox Cole-Green	BCCG	$(0, \infty)$	ident.	log	ident.	-
Box-Cox Cole-Green orig.	BCCGo	$(0, \infty)$	log	log	ident.	-
Box-Cox power exponential	BCPE	$(0, \infty)$	ident.	log	ident.	log
Box-Cox power expon. orig.	BCPEo	$(0, \infty)$	log	log	ident.	log
Box-Cox t	BCT	$(0, \infty)$	ident.	log	ident.	log
Box-Cox t orig.	BCTo	$(0, \infty)$	log	log	ident.	log
exponential	EXP	$(0, \infty)$	log	-	-	-
exponential Gaussian	exGAUS	$(-\infty, \infty)$	ident.	log	log	-
exponential gen. beta 2	EGB2()	$(-\infty, \infty)$	ident.	log	log	log
gamma	GA	$(0, \infty)$	log	log	-	-
generalized beta type 1	GB1	$(0, 1)$	logit	logit	log	log
generalized beta type 2	GB2	$(0, \infty)$	log	log	log	log
generalized gamma	GG	$(0, \infty)$	log	log	ident.	-
generalized inv. Gaussian	GIG	$(0, \infty)$	log	log	ident.	-
generalized t	GT	$(-\infty, \infty)$	ident.	log	log	log
Gumbel	GU	$(-\infty, \infty)$	ident.	log	-	-
inverse Gamma	IGAMMA	$(0, \infty)$	log	log	-	-
inverse Gaussian	IG	$(0, \infty)$	log	log	-	-
Johnson's SU repar.	JSU	$(-\infty, \infty)$	ident.	log	ident.	log
Johnson's SU original	JSUo	$(-\infty, \infty)$	ident.	log	ident.	log
logistic	LO	$(-\infty, \infty)$	ident.	log	-	-
logit normal	LOGITNO	$(0, 1)$	ident.	log	-	-
log normal	LOGNO	$(0, \infty)$	ident.	log	-	-

μ : Promedio

σ : Varianza

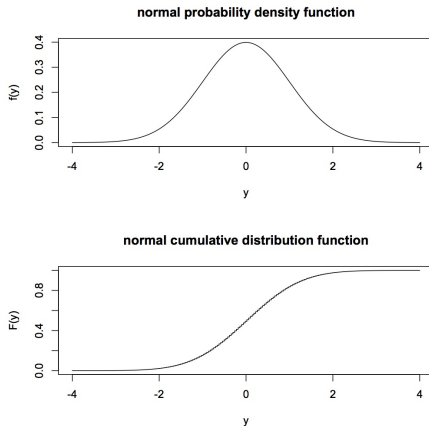
ν : Asimetría

τ : Kurtosis

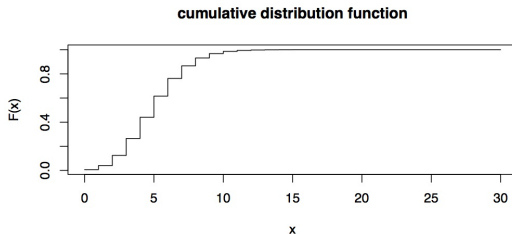
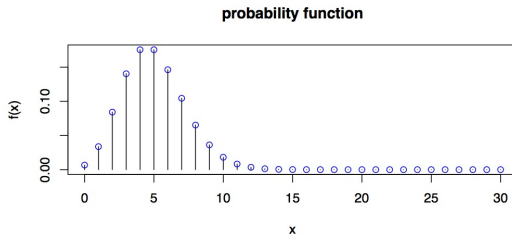
Ejemplo: Una variable que tiene forma de distribución beta

Puede describirse con los dos Primeros parámetros. En cambio, Una exponencial Gausiana se Describe con tres parámetros.

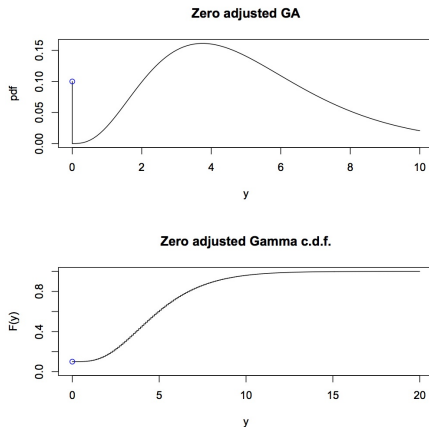
Una de las distribuciones continuas



Una de las distribuciones discretas



Una de las distribuciones mixtas

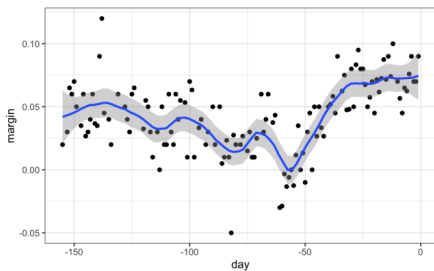


GAMLSS surge como una alternativa que intenta mantener la amplia aplicabilidad de la regresión lineal múltiple, sin estar afectada por los **supuestos del modelo lineal**:

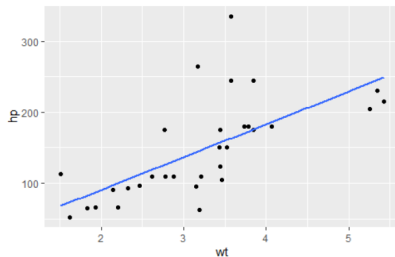
- La relación entre variables es de tipo lineal.
- La homoscedasticidad: La varianza en el error de predicción de la variable dependiente se mantiene constante para cualquier valor de variable independiente.
- La ausencia de Endogeneidad: La correlación entre los residuales y las variables predictoras es cero. En otras palabras, la medición de las variables predictoras no tiene error.
- La distribución subyacente a los residuales y a la variable dependiente es de tipo normal.

Regresión GAMLSS

Ejemplo Regresión No-Lineal



Ejemplo Regresión Lineal



Ejemplo: El Efecto Dunning-Kruger

El efecto Dunning-Kruger ocurre cuando los individuos incompetentes tienden a sobreestimar su habilidad, mientras que los individuos altamente competentes tienden a subestimar su habilidad en relación con la de otros.



Ejemplo: El Efecto Dunning-Kruger

Intelligence 80 (2020) 101449



Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell



The Dunning-Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data



Gilles E. Gignac^{a,*}, Marcin Zajenkowski^b

^a University of Western Australia, Australia

^b Faculty of Psychology, University of Warsaw, Poland

<https://github.com/jcorrean/Pantaleon/blob/main/PantaleonGAMLSS.R>

Ejemplo: El Efecto Dunning-Kruger

```
1 dat<-read.csv(file.choose())
2 # load packages needed
3 library("gamlss")
4 library("gamlss.dist")
5
6 # Modelo de parametros de ubicación, escala y forma
7 muform <- rendimiento ~ pb(TAH) # parametro de ubicación
8 sigmaform <- tauform <- nuform <- ~ pb(TAH) # parametro de escala
9
10
11 m1 <- gamlss(muform, family=NO, data=na.omit(dat)) # modelo homoscedastico
12 m2 <- gamlss(muform, sigmaform, family=NO, data=na.omit(dat)) # modelo heteroscedastico
```

Ejemplo: El Efecto Dunning-Kruger

```
15 AIC(m1,m2)
16 # Observe otras formas de modelar la relación entre
17 # variables, partiendo de un modelo con dos parametros
18
19 m3 <-gamlss(muform, sigmaform, family=L0, data=na.omit(dat))
20 m4 <-gamlss(muform, sigmaform, family=GU, data=na.omit(dat))
21 m5 <-gamlss(muform, sigmaform, family=RG, data=na.omit(dat))
22
23 # Ahora veamos modelos con distribuciones que requieran
24 # tres parametros,
25 m7 <-gamlss(muform, sigmaform, tauform, family=TF, data=na.omit(dat))
26 m8 <-gamlss(muform, sigmaform, tauform, family=SN1, data=na.omit(dat))
27
28 # Y ahora veamos otras formas de distribución con
29 # cuatro parametros
30 m22 <-gamlss(muform, sigmaform, tauform, nuform, family=NET, data=na.omit(dat))
31
```

Ejemplo: El Efecto Dunning-Kruger

```
32 AIC(m1, m2, m3, m4, m5, m7, m22)
33 #
34 #           df           AIC
35 # m2  4.128266 10732.88
36 # m1  3.181336 10733.73
37 # m7  8.875680 10737.76
38 # m3  4.001297 10761.08
39 # m22 4.000869 10843.39
40 # m5  5.034416 10897.78
41 # m4  4.001317 10992.14
```

El criterio de información de Akaike (AIC) es una medida de la calidad relativa de un modelo estadístico, para un conjunto dado de datos. AIC proporciona un medio para la selección del modelo.