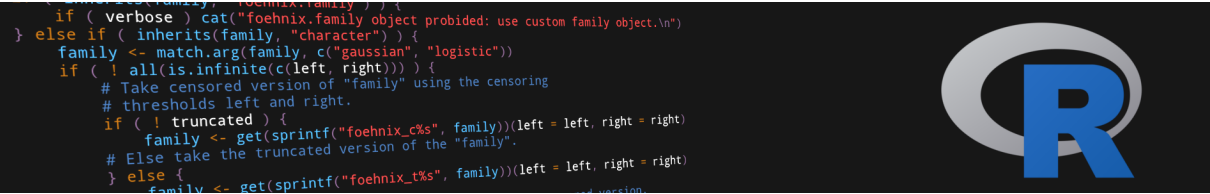
universität innsbruck

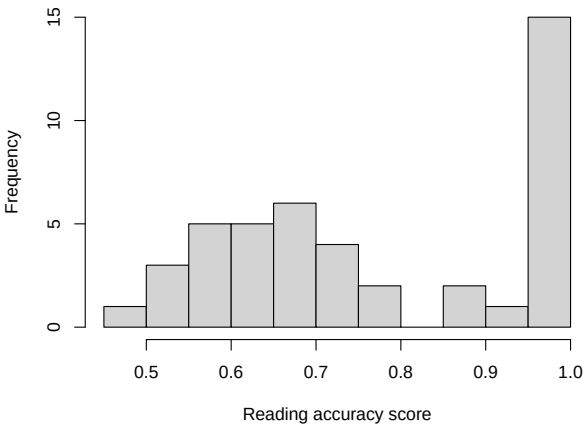


Regression Models for $[0, 1]$ Responses Using betareg and crch

Ioannis Kosmidis, Achim Zeileis

https://www.zeileis.org/

Motivation

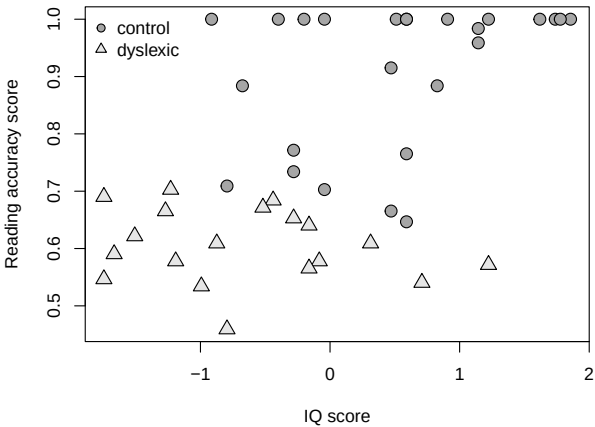


Goal: Model limited response variables in unit interval.

Examples: Fractions or proportions (not from independent Bernoulli trials).

Illustration: Reading `accuracy` of 44 primary school children, explained by `dyslexia` status and `iq` score.

Motivation

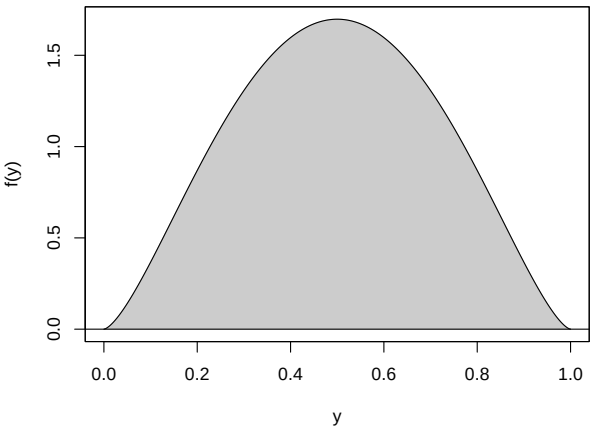


Goal: Model limited response variables in unit interval.

Examples: Fractions or proportions (not from independent Bernoulli trials).

Illustration: Reading `accuracy` of 44 primary school children, explained by `dyslexia` status and `iq` score.

Beta distribution



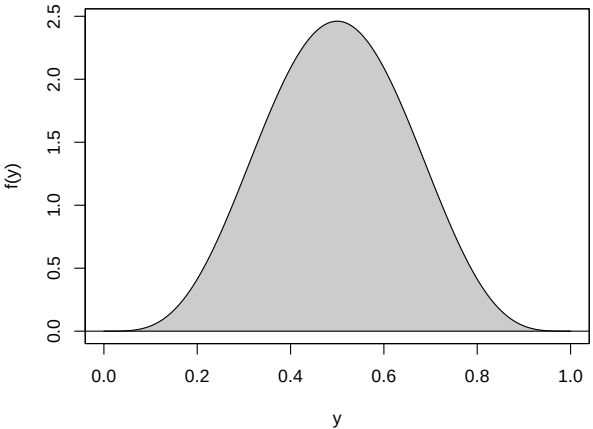
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



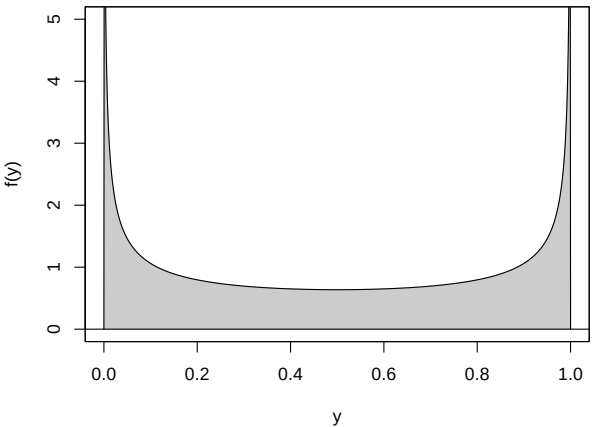
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



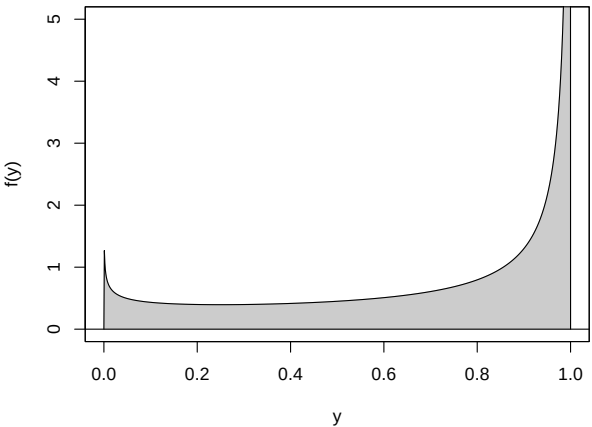
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



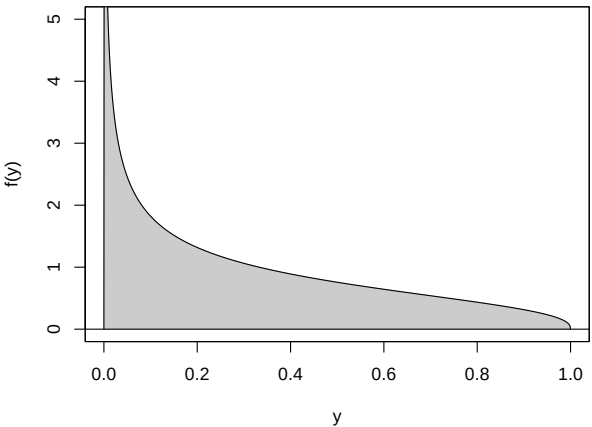
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



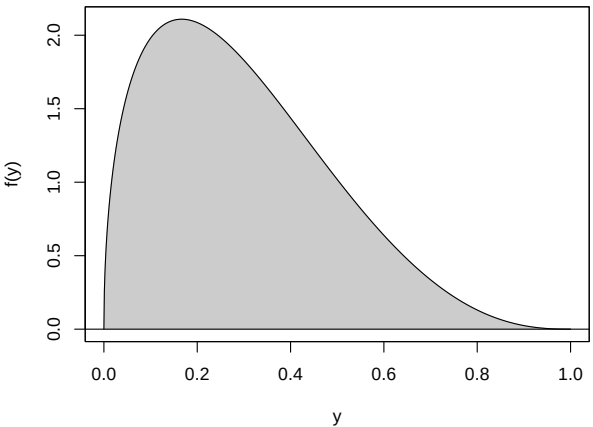
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



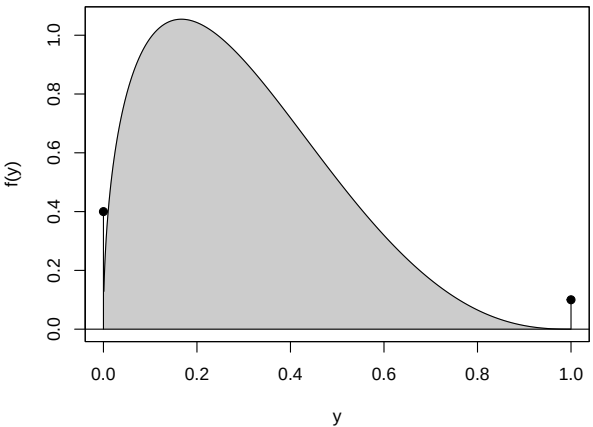
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Zero-and/or-one-inflated beta distribution



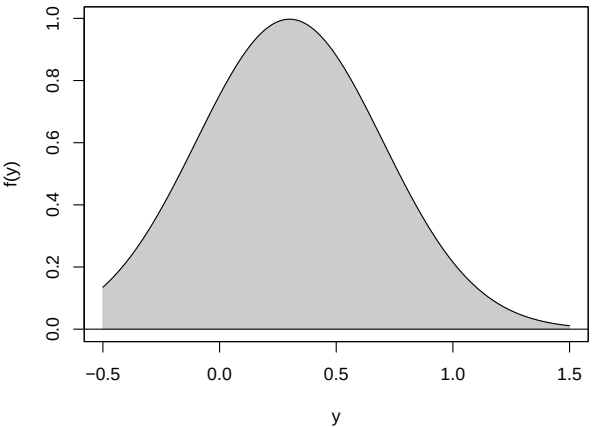
Parameters: Mean μ, precision ϕ, point masses π_0, π_1.

Regression: Link all four parameters to predictors.

Advantage: Keep flexibility, accomodate boundaries.

Disadvantage: Many parameters, separate determinants for boundaries.

Censored normal distribution (tobit)



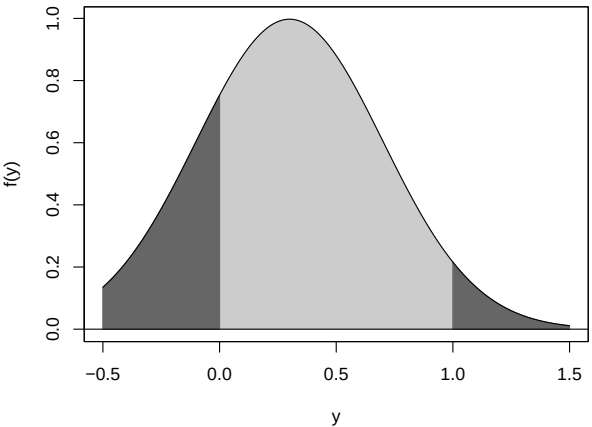
Parameters: Mean μ, variance σ^2.

Regression: Link both parameters to predictors.

Advantage: No additional determinants for boundaries.

Disadvantage: Less flexible than beta.

Censored normal distribution (tobit)



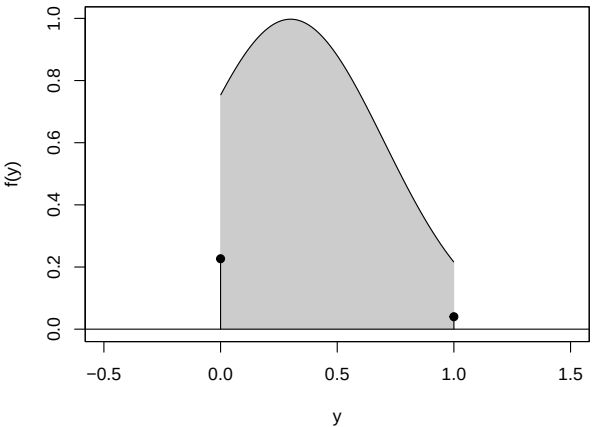
Parameters: Mean μ, variance σ^2.

Regression: Link both parameters to predictors.

Advantage: No additional determinants for boundaries.

Disadvantage: Less flexible than beta.

Censored normal distribution (tobit)



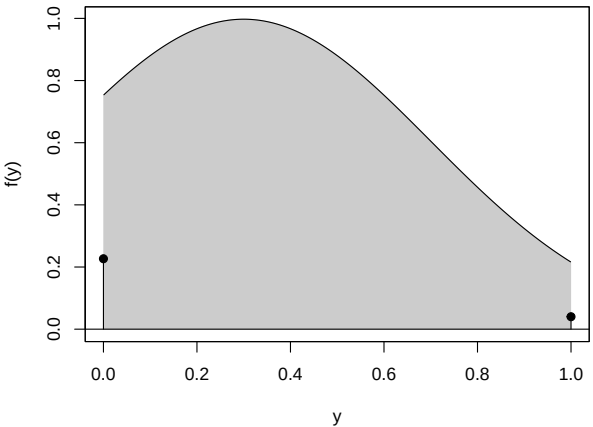
Parameters: Mean μ, variance σ^2.

Regression: Link both parameters to predictors.

Advantage: No additional determinants for boundaries.

Disadvantage: Less flexible than beta.

Censored normal distribution (tobit)



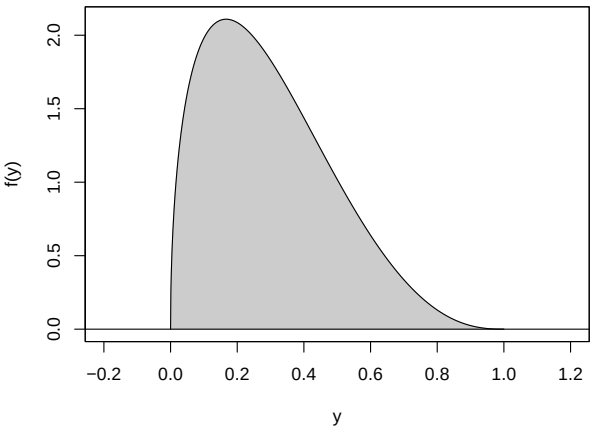
Parameters: Mean μ, variance σ^2.

Regression: Link both parameters to predictors.

Advantage: No additional determinants for boundaries.

Disadvantage: Less flexible than beta.

Extended-support beta mixture distribution



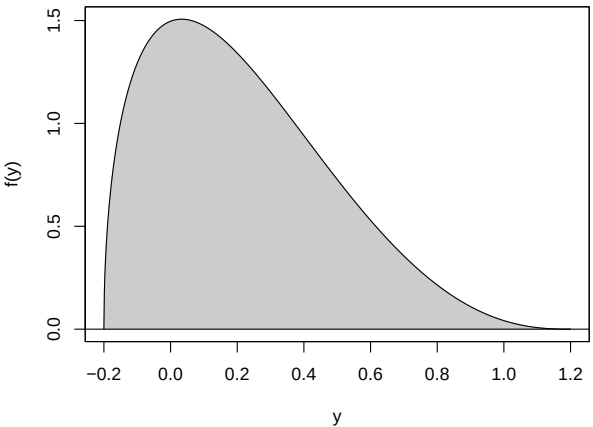
Parameters: Mean μ, precision ϕ, exceedence ν.

Regression: Link only μ and ϕ to predictors.

Advantage: Single parameter ν links normal and beta.

Disadvantage: Somewhat more complex.

Extended-support beta mixture distribution



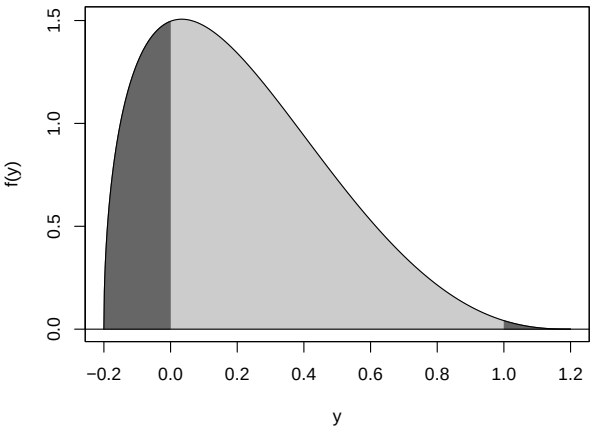
Parameters: Mean μ, precision ϕ, exceedence ν.

Regression: Link only μ and ϕ to predictors.

Advantage: Single parameter ν links normal and beta.

Disadvantage: Somewhat more complex.

Extended-support beta mixture distribution



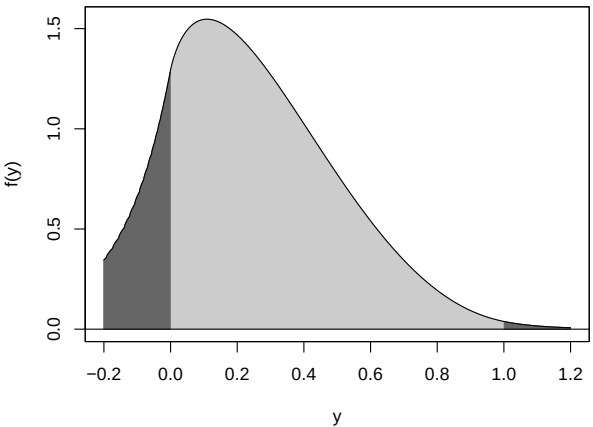
Parameters: Mean μ, precision ϕ, exceedence ν.

Regression: Link only μ and ϕ to predictors.

Advantage: Single parameter ν links normal and beta.

Disadvantage: Somewhat more complex.

Extended-support beta mixture distribution



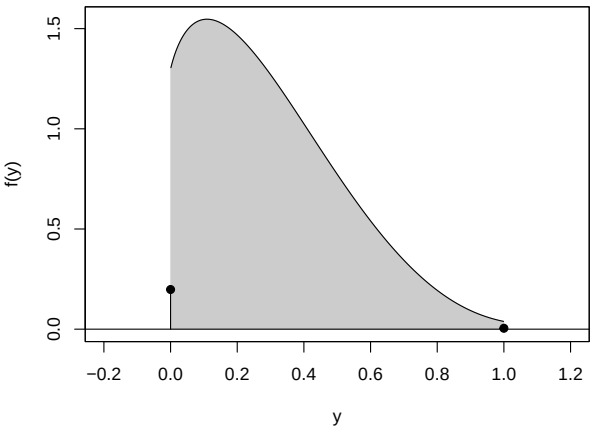
Parameters: Mean μ, precision ϕ, exceedence ν.

Regression: Link only μ and ϕ to predictors.

Advantage: Single parameter ν links normal and beta.

Disadvantage: Somewhat more complex.

Extended-support beta mixture distribution



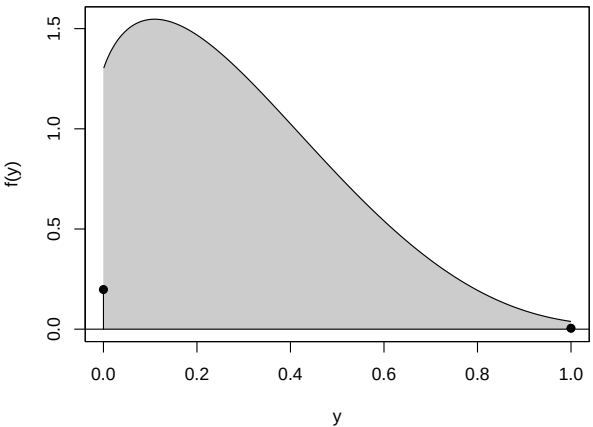
Parameters: Mean μ, precision ϕ, exceedence ν.

Regression: Link only μ and ϕ to predictors.

Advantage: Single parameter ν links normal and beta.

Disadvantage: Somewhat more complex.

Extended-support beta mixture distribution



Parameters: Mean μ, precision ϕ, exceedence ν.

Regression: Link only μ and ϕ to predictors.

Advantage: Single parameter ν links normal and beta.

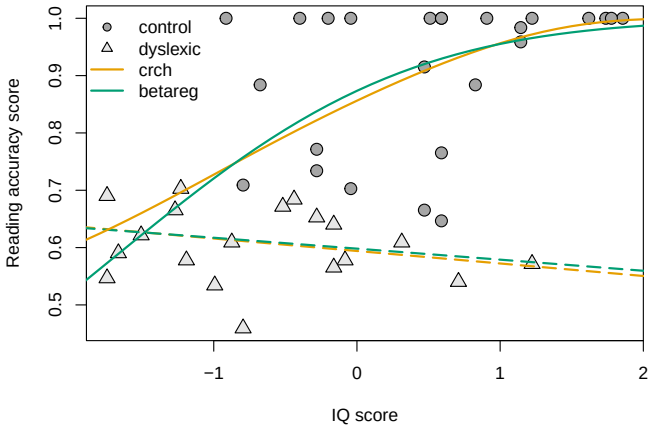
Disadvantage: Somewhat more complex.

R packages

```
R> crch(accuracy1 ~ dyslexia * iq | dyslexia + iq, data = ReadingSkills,
+    left = 0, right = 1)
R> betareg(accuracy1 ~ dyslexia * iq | dyslexia + iq, data = ReadingSkills)
```

R packages

```
R> crch(accuracy1 ~ dyslexia * iq | dyslexia + iq, data = ReadingSkills,
+    left = 0, right = 1)
R> betareg(accuracy1 ~ dyslexia * iq | dyslexia + iq, data = ReadingSkills)
```



References

Cribari-Neto F, Zeileis A (2010). “Beta Regression in R.” *Journal of Statistical Software*, **34**(2), 1–24. doi:10.18637/jss.v034.i02

Messner JW, Mayr GJ, Zeileis A (2016). “Heteroscedastic Censored and Truncated Regression with crch.” *The R Journal*, **8**(1), 173–181. doi:10.32614/RJ-2016-012

Kosmidis I, Zeileis A (2024). “Extended-Support Beta Regression for $[0, 1]$ Responses.” *arXiv.org E-Print Archive*, Forthcoming.

Software:
https://CRAN.R-project.org/package=crch (Version 1.1-2)
https://CRAN.R-project.org/package=betareg (Version 3.2-0)

Mastodon: @zeileis@fosstodon.org

X: @AchimZeileis

Web: https://www.zeileis.org/