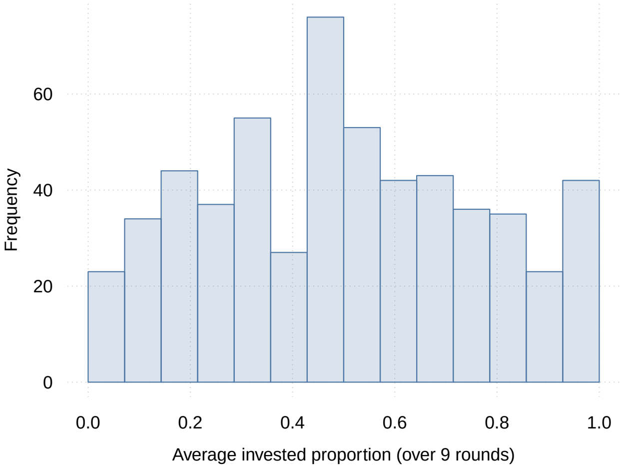
universität innsbruck

Extended-Support Beta Regression for $[0, 1]$ Responses

Ioannis Kosmidis, Achim Zeileis

https://www.zeileis.org/

Motivation

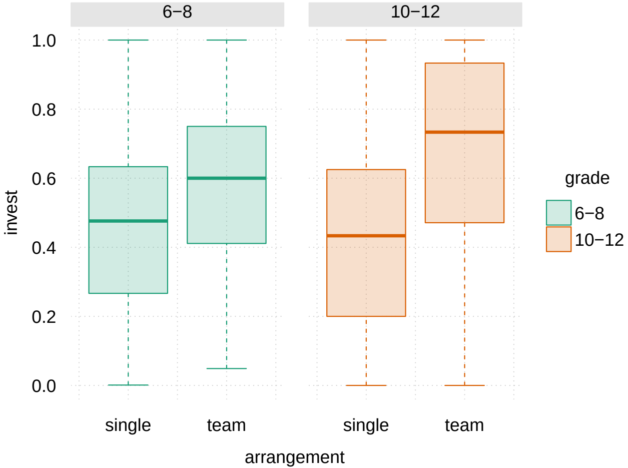


Goal: Model limited response variables in unit interval.

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

Illustration: `investment` in a risky lottery with positive expected payout, explained by `arrangement`, `grade`, ... of high-school students.

Motivation

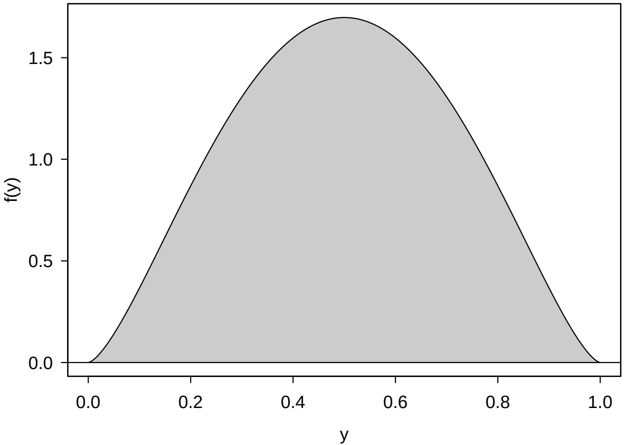


Goal: Model limited response variables in unit interval.

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

Illustration: `investment` in a risky lottery with positive expected payout, explained by `arrangement`, `grade`, . . . of high-school students.

Beta distribution



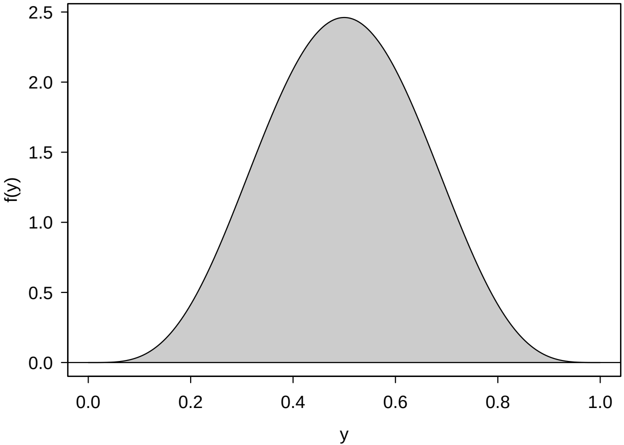
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



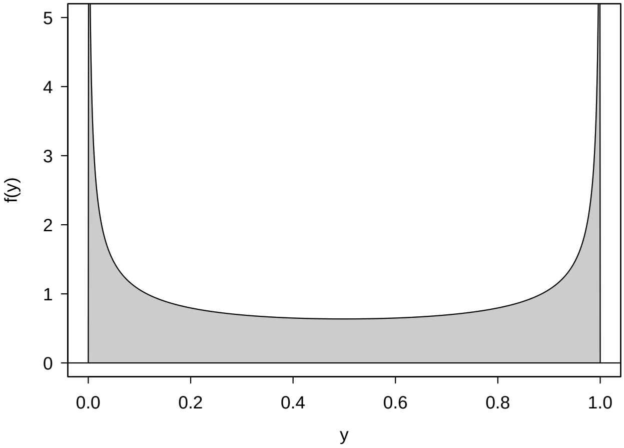
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



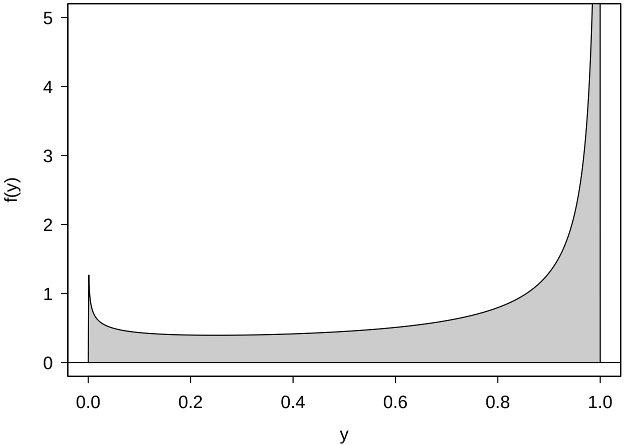
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



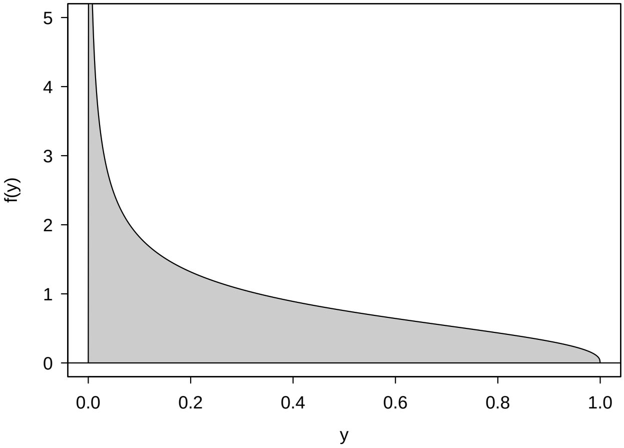
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



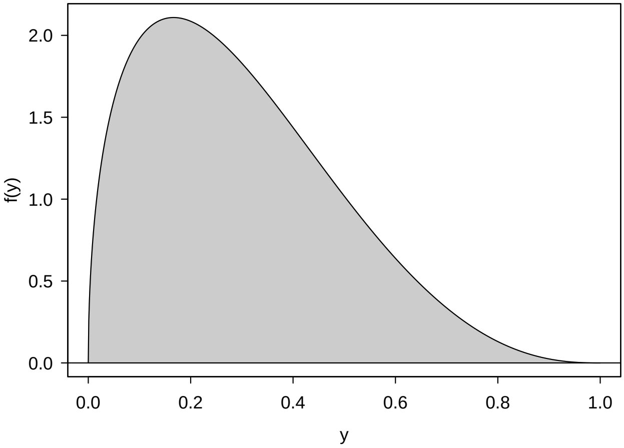
Parameters: Mean μ, precision ϕ.

Regression: Link both parameters to predictors.

Advantage: Flexible shape, full likelihood.

Disadvantage: Zero probability for 0 and 1.

Beta distribution



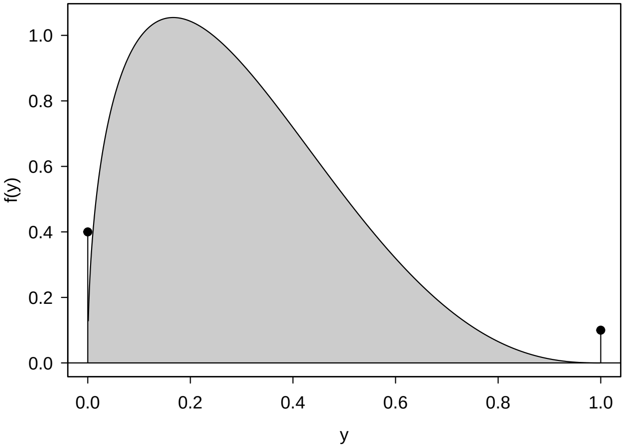
Parameters: Mean μ, precision ϕ.

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



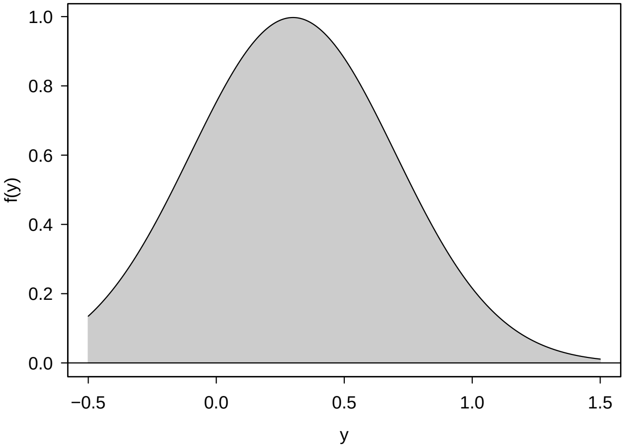
Parameters: Mean μ, precision ϕ, point masses π_0, π_1.

Regression: Link all four parameters to predictors.

Advantage: Keep flexibility, accomodate boundaries.

Disadvantage: Many parameters, separate determinants for boundaries.

Censored normal distribution (tobit)



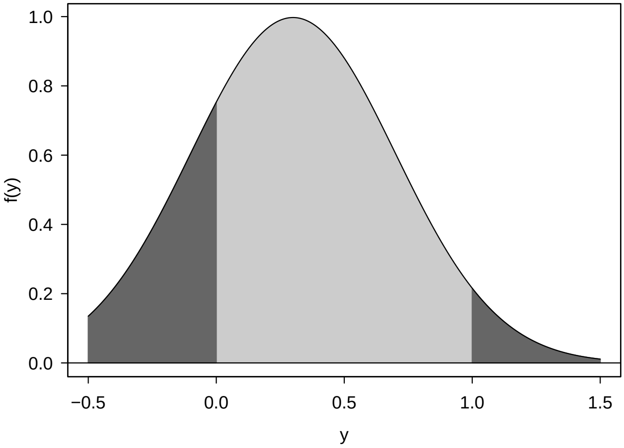
Parameters: Mean μ, variance σ^2.

Regression: Link both parameters to predictors.

Advantage: No additional determinants for boundaries.

Disadvantage: Less flexible than beta.

Censored normal distribution (tobit)



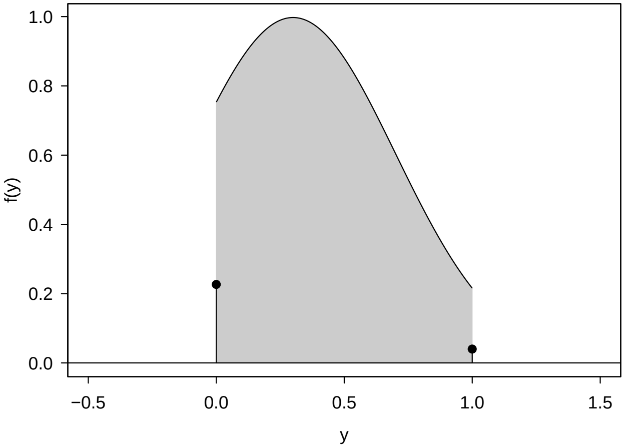
Parameters: Mean μ, variance σ^2.

Regression: Link both parameters to predictors.

Advantage: No additional determinants for boundaries.

Disadvantage: Less flexible than beta.

Censored normal distribution (tobit)



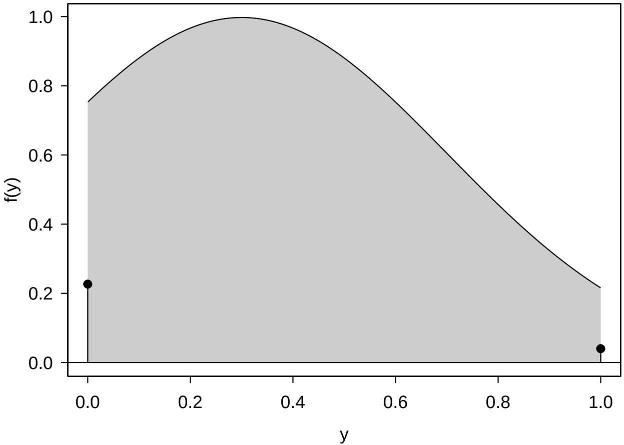
Parameters: Mean μ, variance σ^2.

Regression: Link both parameters to predictors.

Advantage: No additional determinants for boundaries.

Disadvantage: Less flexible than beta.

Censored normal distribution (tobit)



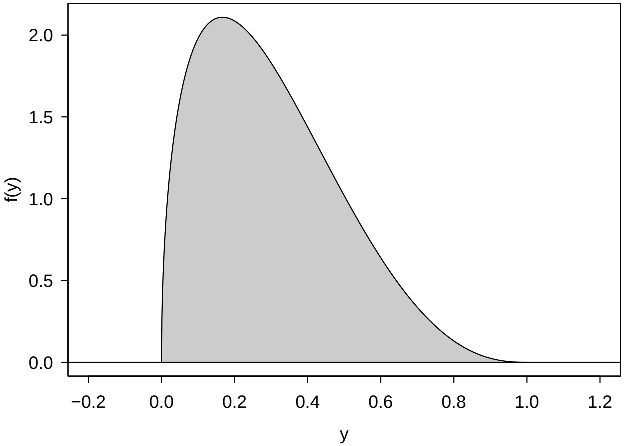
Parameters: Mean μ, variance σ^2.

Regression: Link both parameters to predictors.

Advantage: No additional determinants for boundaries.

Disadvantage: Less flexible than beta.

Extended-support beta mixture distribution



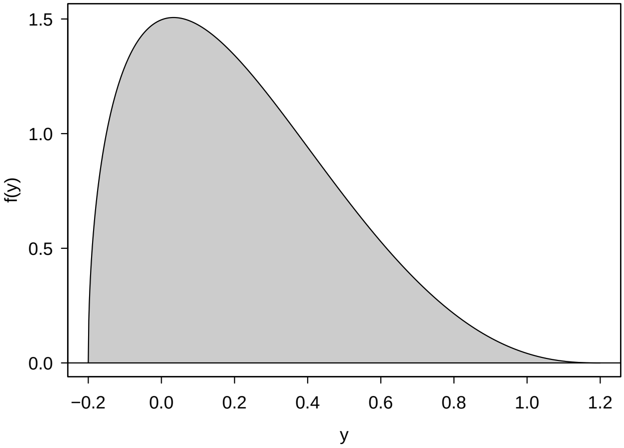
Parameters: Mean μ, precision ϕ, exceedence ν.

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



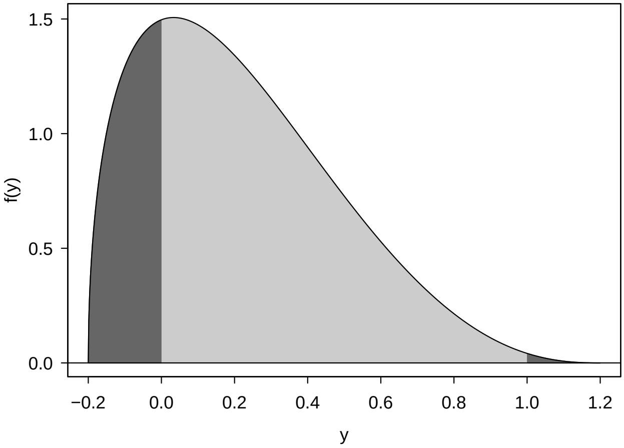
Parameters: Mean μ, precision ϕ, exceedence ν.

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



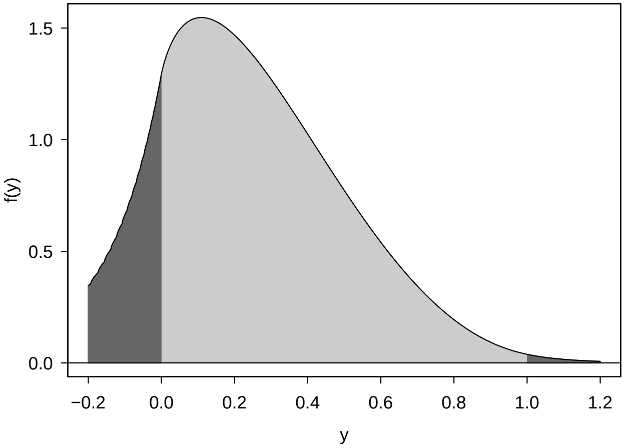
Parameters: Mean μ, precision ϕ, exceedence ν.

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



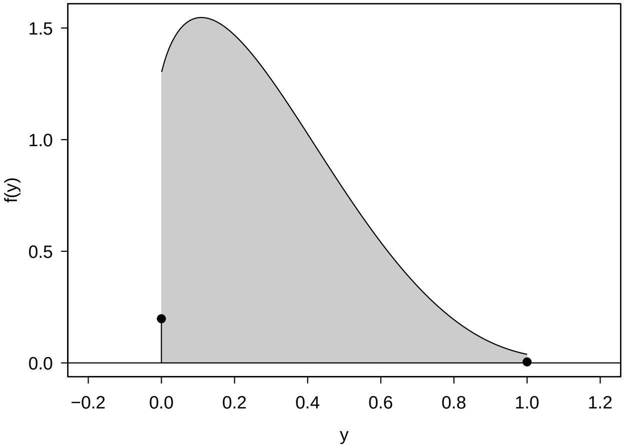
Parameters: Mean μ, precision ϕ, exceedence ν.

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



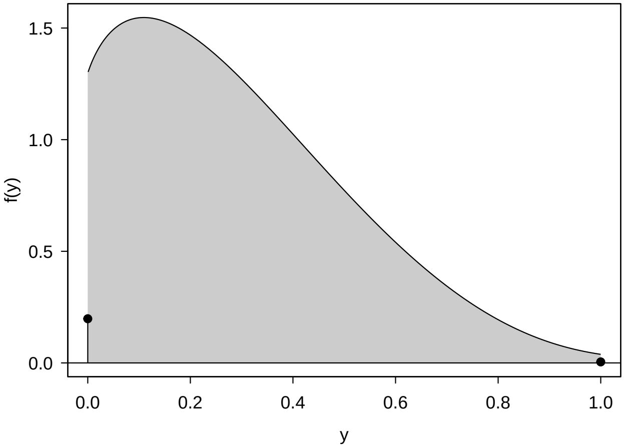
Parameters: Mean μ, precision ϕ, exceedence ν.

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.

Theory

More formally: $\mathrm{XB}(\mu, \phi, u)$ is a beta distribution $\mathrm{B}(\mu, \phi)$ with support extended to $(-u, 1 + u)$ and censored at 0 and 1.

Theory

More formally: $\mathrm{XB}(\mu, \phi, u)$ is a beta distribution $\mathrm{B}(\mu, \phi)$ with support extended to $(-u, 1+u)$ and censored at 0 and 1.

$$f_{(\mathrm{XB})}(y \mid \mu, \phi, u) \quad = \quad f_{(\mathrm{B})}\left(\frac{y+u}{1+2u} \,\middle|\, \mu, \phi\right) \frac{1}{1+2u}\,, \quad \text{if } y \in (0, 1)$$

Theory

More formally: $\text{XB}(\mu, \phi, u)$ is a beta distribution $\text{B}(\mu, \phi)$ with support extended to $(-u, 1+u)$ and censored at 0 and 1.

$$f_{(\text{XB})}(y \mid \mu, \phi, u) \quad = \quad \begin{cases} F_{(\text{B})}\left(\frac{u}{1+2u} \,\middle|\, \mu, \phi\right), & \text{if } y = 0 \\ f_{(\text{B})}\left(\frac{y+u}{1+2u} \,\middle|\, \mu, \phi\right)\frac{1}{1+2u}, & \text{if } y \in (0, 1) \\ 1 - F_{(\text{B})}\left(\frac{1+u}{1+2u} \,\middle|\, \mu, \phi\right), & \text{if } y = 1 \end{cases}$$

Theory

More formally: $\mathrm{XB}(\mu, \phi, u)$ is a beta distribution $\mathrm{B}(\mu, \phi)$ with support extended to $(-u, 1+u)$ and censored at 0 and 1.

$$f_{(\mathrm{XB})}(y \mid \mu, \phi, u) \quad = \quad \begin{cases} F_{(\mathrm{B})}\left(\dfrac{u}{1+2u} \;\middle|\; \mu, \phi\right), & \text{if } y = 0 \\ f_{(\mathrm{B})}\left(\dfrac{y+u}{1+2u} \;\middle|\; \mu, \phi\right) \dfrac{1}{1+2u}, & \text{if } y \in (0,1) \\ 1 - F_{(\mathrm{B})}\left(\dfrac{1+u}{1+2u} \;\middle|\; \mu, \phi\right), & \text{if } y = 1 \end{cases}$$

Special cases: Beta ($u = 0$) and censored normal ($u \to \infty$) distributions.

Theory

Shrinkage: $\text{XBX}(\mu, \phi, \nu)$ is a continuous mixture of $\text{XB}(\mu, \phi, u)$ with $u \sim \text{Exp}(\nu)$.

Theory

Shrinkage: XBX(μ, ϕ, ν) is a continuous mixture of XB(μ, ϕ, u) with $u \sim \text{Exp}(\nu)$.

$$f_{(\text{XBX})}(y \mid \mu, \phi, \nu) \quad = \quad \nu^{-1} \int_0^\infty f_{(\text{XB})}(y \mid \mu, \phi, u) \, e^{-u/\nu} du$$

Loss aversion

Behavioral economics experiment: Glätzle-Rützler *at al.* (2015).

- Determinants of loss aversion in high-school students.
- Proportion of tokens invested in risky lottery with positive expected payouts.
- Outcome: Average investments over nine rounds.
- Experimental factors: `grade` (lower vs. upper), `arrangement` (single vs. team of two), `male` (at least one), and `age`.

Loss aversion

Behavioral economics experiment: Glätzle-Rützler *at al.* (2015).

- Determinants of loss aversion in high-school students.
- Proportion of tokens invested in risky lottery with positive expected payouts.
- Outcome: Average investments over nine rounds.
- Experimental factors: `grade` (lower vs. upper), `arrangement` (single vs. team of two), `male` (at least one), and `age`.

Of interest: Extent of risk aversion.

- Mean investments: $E(Y)$.
- Probability to behave like a rational *homo oeconomicus*: $P(Y > 0.95)$.

Loss aversion

Original analysis: Linear regression model for mean only.

```
R> la_ols <- glm(invest ~ grade * (arrangement + age) + male, data = LossAversion)
```

Loss aversion

Original analysis: Linear regression model for mean only.

```
R> la_ols <- glm(invest ~ grade * (arrangement + age) + male, data = LossAversion)
```

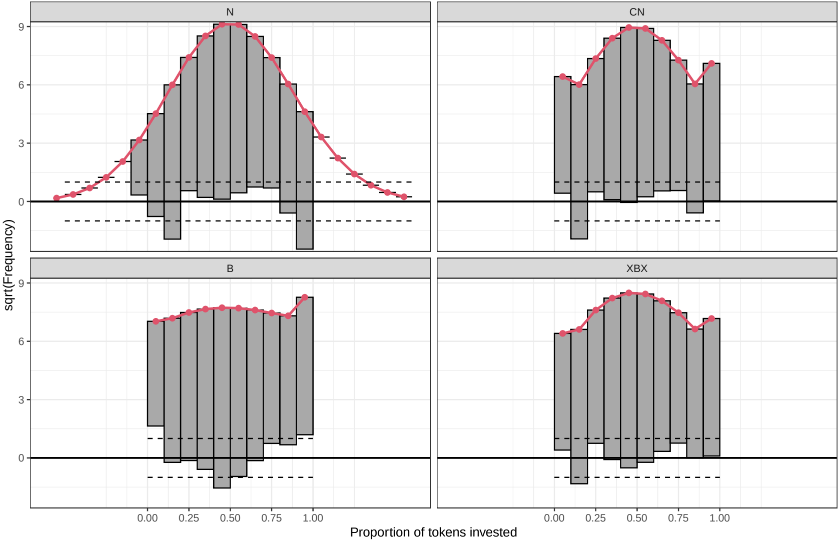
Alternatively: Probabilistic models to simultaneously model mean and probability.

- N: Linear regression, interpreted as homoscedastic normal model.
- CN: Heteroscedastic censored normal model.
- B: Beta regression after ad-hoc scaling to the open unit interval.
- XBX: Extended-support beta mixture model.

```
R> la_xbx <- betareg(invest ~ grade * (arrangement + age) + male |
+    arrangement + male + grade, data = LossAversion)
```

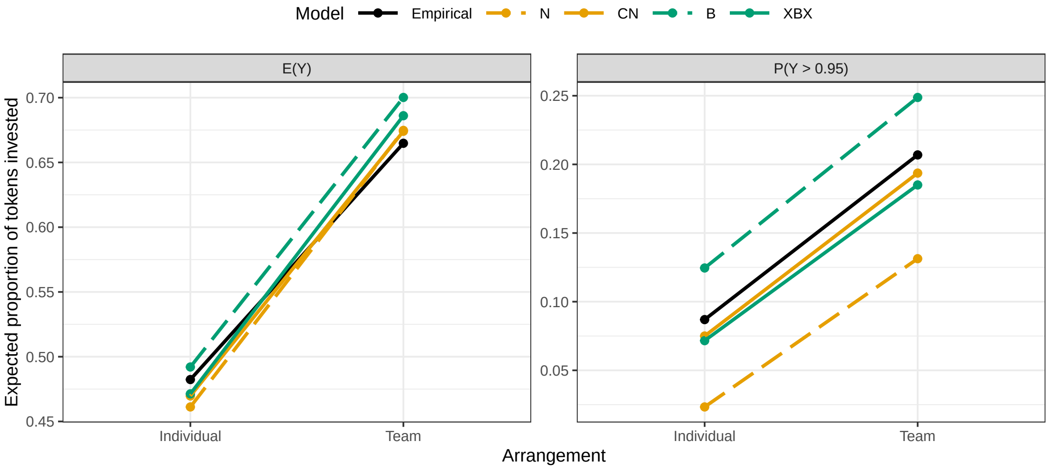
etc.

Loss aversion



Loss aversion

Arrangement effects: For 16-year old, (at least one) male players.



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

Kosmidis I, Zeileis A (2024). "Extended-Support Beta Regression for $[0, 1]$ Responses." *arXiv.org E-Print Archive*, arXiv:2409.07233. doi:10.48550/arXiv.2409.07233

Software:
https://topmodels.R-Forge.R-project.org/betareg/

Mastodon: @zeileis@fosstodon.org

Bluesky: @zeileis.org

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