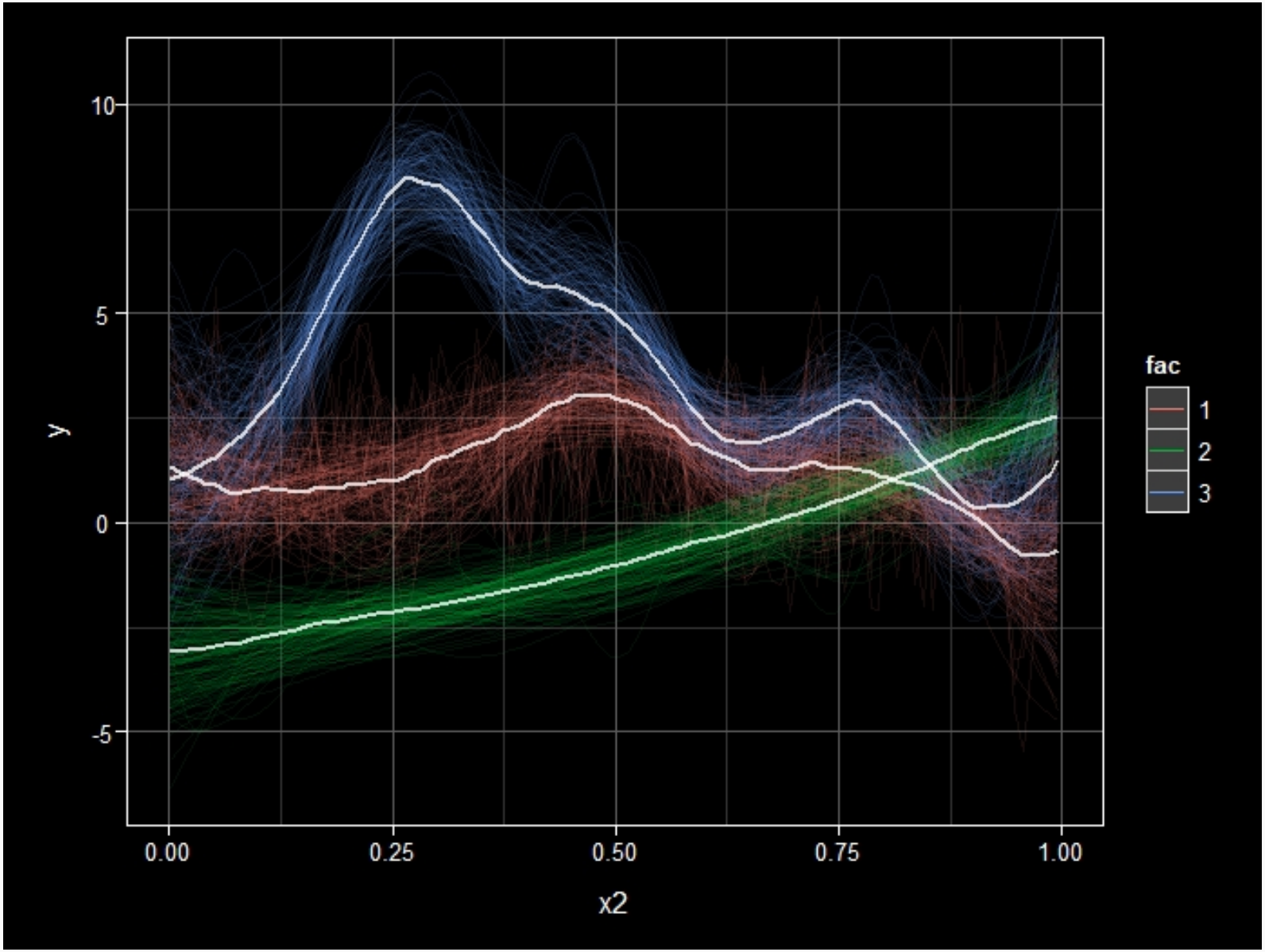
Custom Response Distributions with brms

Paul Bürkner

2023-12-22

Idea of brms



Response Distributions (*Families*) in brms

Play a central role in the likehood $p(y|\theta)$

Examples:

- Gaussian
- Poisson
- Gamma
- ...

Misspecifying the family may seriously distort your results

brms natively supports 37 families to date

FAQ: Can I use brms with other families?

Case Study: Housing Rents in Munich

```
data("rent99", package = "gamlss.data")
```

rentsqm	area	yearc	district
4.228797	26	1918	916
8.688646	28	1918	813
8.721369	30	1918	611
3.547009	30	1918	2025
4.446154	30	1918	561
11.300851	30	1918	541
6.942928	31	1918	822
10.426800	31	1918	1713
6.759615	32	1918	1812
7.432790	33	1918	152

Housing Rents in Munich: Gamma Model

Density of the Gamma distribution:

$$p(y|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} \exp(-\beta y)$$

with mean and variance:

- $\mathsf{E}(y) = \mu = \frac{\alpha}{\beta}$
- $\mathsf{Var}(y) = v = \frac{\alpha}{\beta^2} = \frac{\mu^2}{\alpha}$

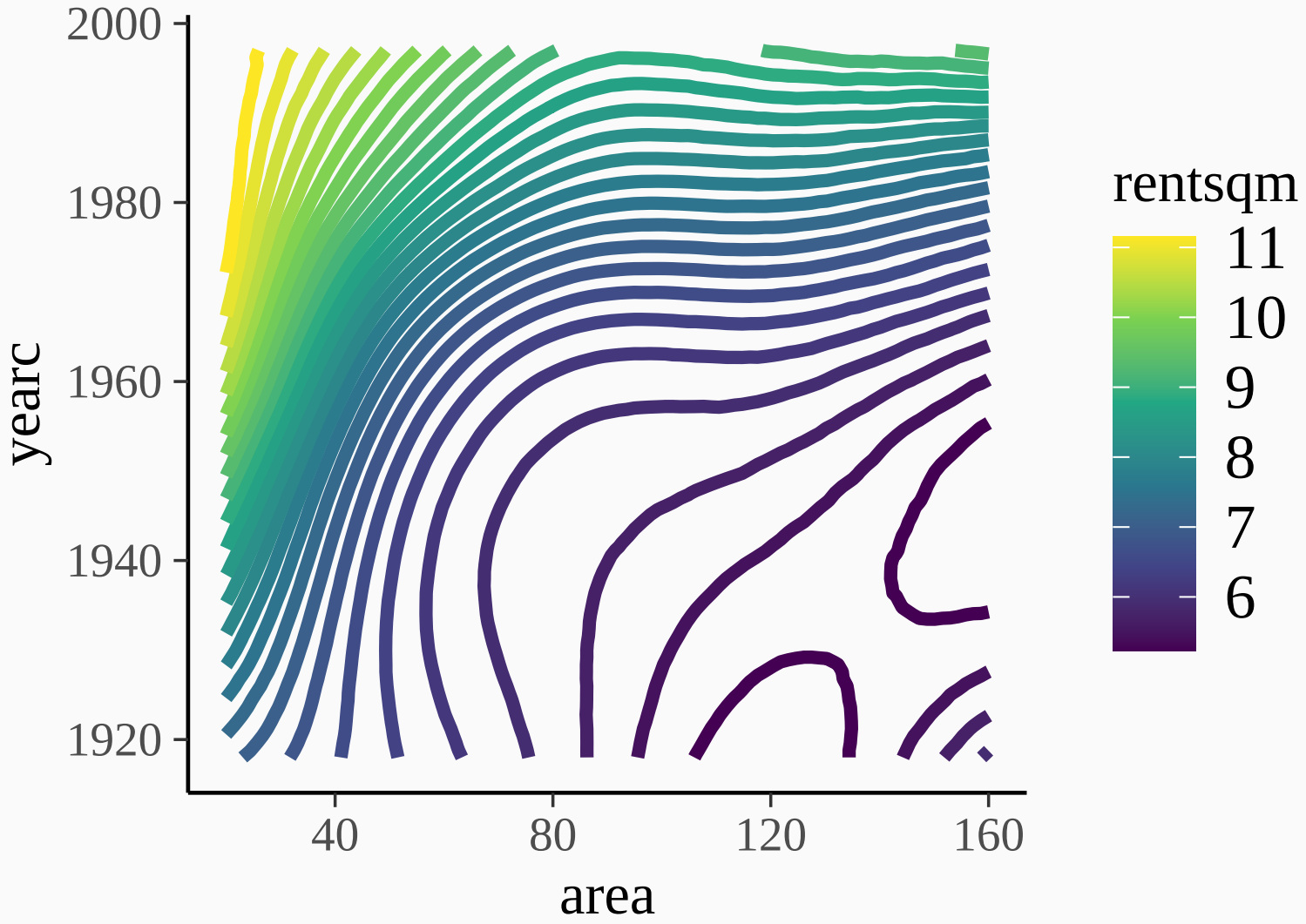
We reparameterize $\beta = \frac{\alpha}{\mu}$ to estimate μ and α

Fit a multilevel Gamma model in brms:

```
fit1 <- brm(
  rentsqm ~ t2(area, yearc) + (1 | district),
  data = rent99, family = Gamma("log")
)
```

Gamma Model: Visualization

```
conditional_effects(fit1, "area:yearc", surface = TRUE)
```



The Mean-Variance Gamma Model

What if we want to predict both the mean and the variance?

We reparameterize $\alpha = \frac{\mu^2}{v}$ and $\beta = \frac{\mu}{v}$

Then we have mean and variance:

- $\mathrm{E}(y) = \mu$
- $\mathrm{Var}(y) = v$

Specifying the log-density in Stan:

```
real gamma2_lpdf(real y, real mu, real v) {
  return gamma_lpdf(y | mu * mu / v, mu / v);
}
```

Custom Families in brms

Define a `custom_family` in brms:

```
gamma2 <- custom_family(
  name = "gamma2",
  dpars = c("mu", "v"),
  links = c("log", "log"),
  type = "real", lb = c(0, 0)
)
```

Prepare the required Stan code:

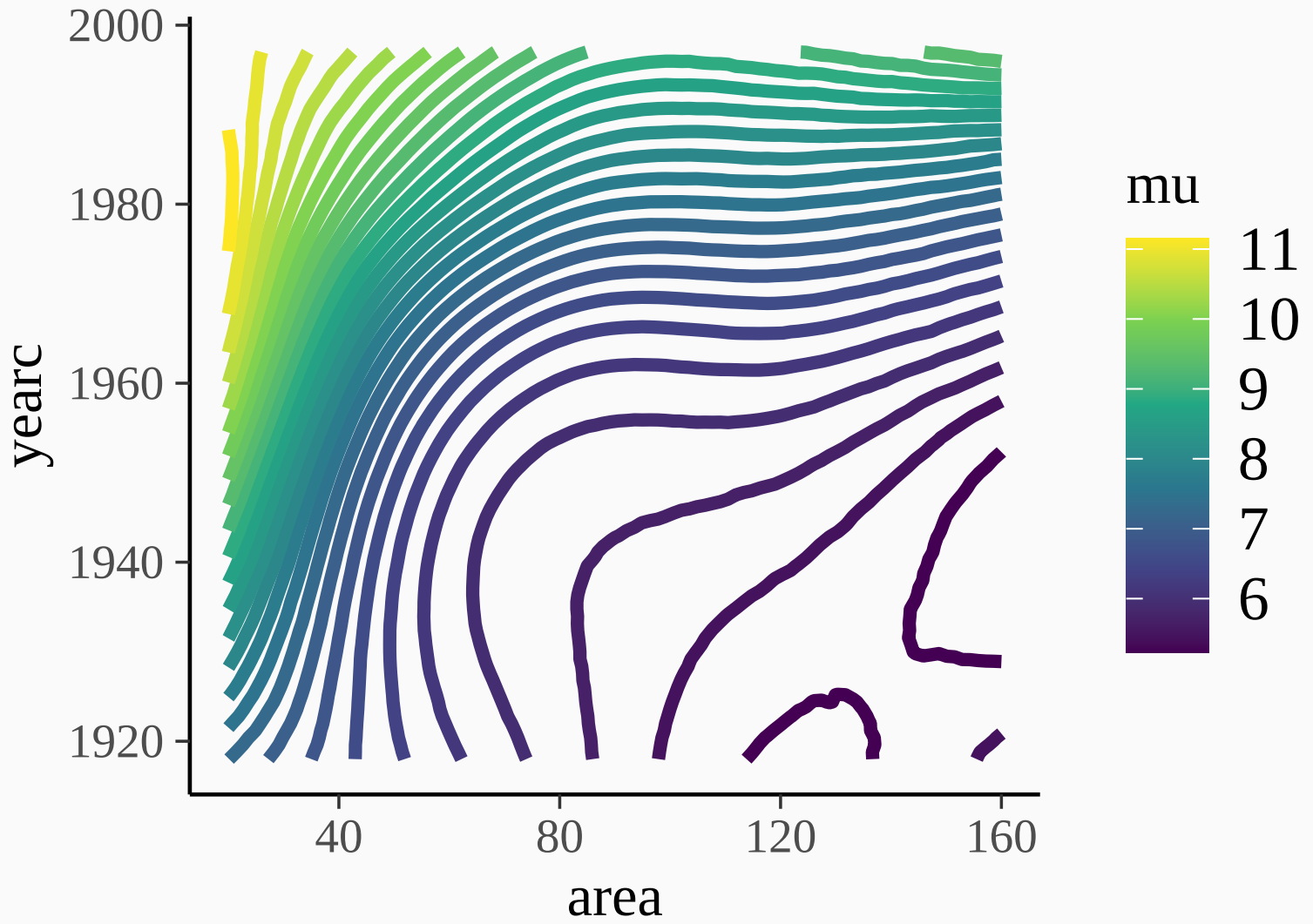
```
stan_gamma2 <- "
  real gamma2_lpdf(real y, real mu, real v) {
    return gamma_lpdf(y | mu * mu / v, mu / v);
  }
"
```

The Mean-Variance Gamma Model: Model Fitting

```
bform2 <- bf(
  rentsqm ~ t2(area, yearc) + (1 | p | district),
  v ~ t2(area, yearc) + (1 | p | district)
)
stanvars2 <- stanvar(
  scode = stan_gamma2, block = "functions"
)
fit2 <- brm(
  bform2, data = rent99,
  family = gamma2,
  stanvars = stanvars2
)
```

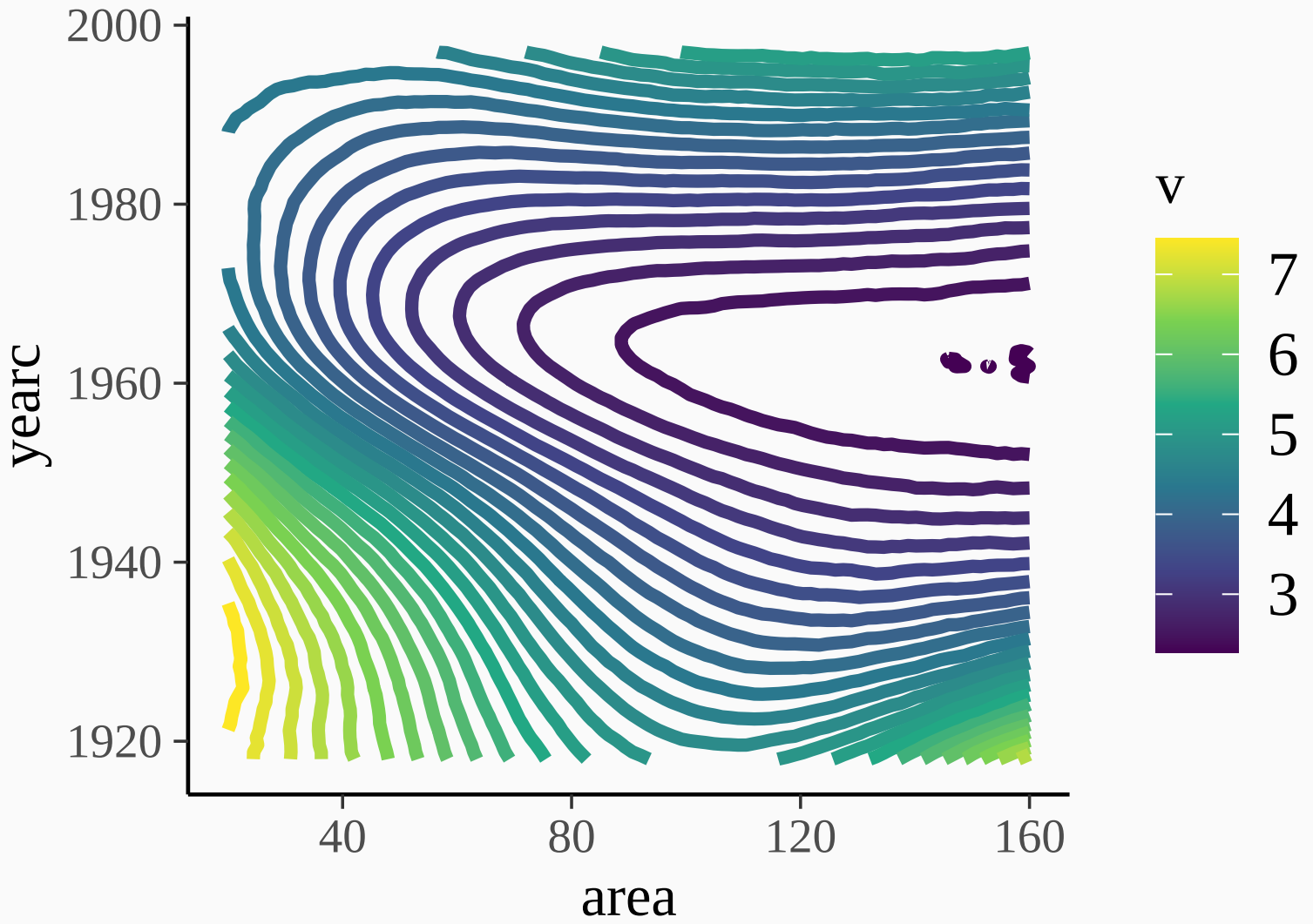
Visualiation of the Mean

```
conditional_effects(fit2, effects = "area:yearc",
                    surface = TRUE, dpar = "mu")
```



Visualiation of the Variance

```
conditional_effects(fit2, effects = "area:yearc",
                    surface = TRUE, dpar = "v")
```



Model comparison via LOO-CV

Define a `log_lik_<family>` function:

```
log_lik_gamma2 <- function(i, draws) {
  # see case study
}
```

```
loo_compare(loo(fit1), loo(fit2))
```

```
##      elpd_diff se_diff
## fit2    0.0       0.0
## fit1 -162.7      23.6
```