Thomas W. Yee

*Figures from "Vector Generalized Linear and Additive Models: With an Implementation in R"

February 29, 2016

Springer

Preface

This document gives the figures from Yee (2015), along with their captions. There are slight differences between these figures and in the book, due to edits by Springer. I retain the copyright for all these figures. They can be used for private study and teaching, but cannot be used in publications without my written consent.

Thomas Yee
Auckland, New Zealand
November 2015

Some, imagining they can best commend themselves to the Eternal by means of statues, are eagerly desirous of them, as if they were certain to obtain more reward from brazen figures unendowed with sense, than from the consciousness of duties honourably and uprightly performed.
—Ammianus Marcellinus

Contents

Chapter 1
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

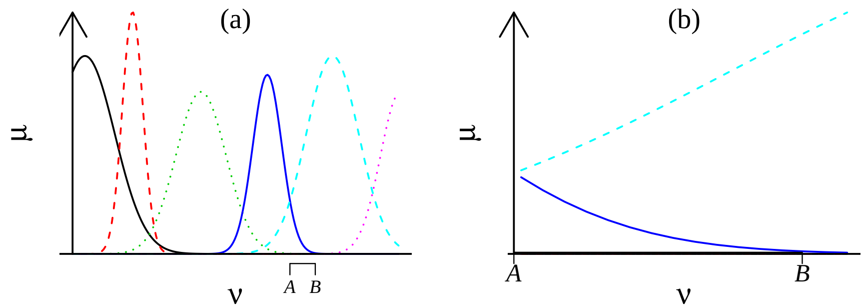


Fig. 1.1 Rank-1 constrained quadratic ordination (CQO). (a) The mean abundance is $\mu_s(\nu) = E(Y_s|\nu)$ for $s = 1, \ldots, S$ species, and $\nu = \boldsymbol{c}^T \boldsymbol{x}_2$ is a latent variable. (b) Zooming in on the subinterval $[A, B]$ in (a). This is approximately linearly on the η scale, meaning a RR-VGLM would be suitable.

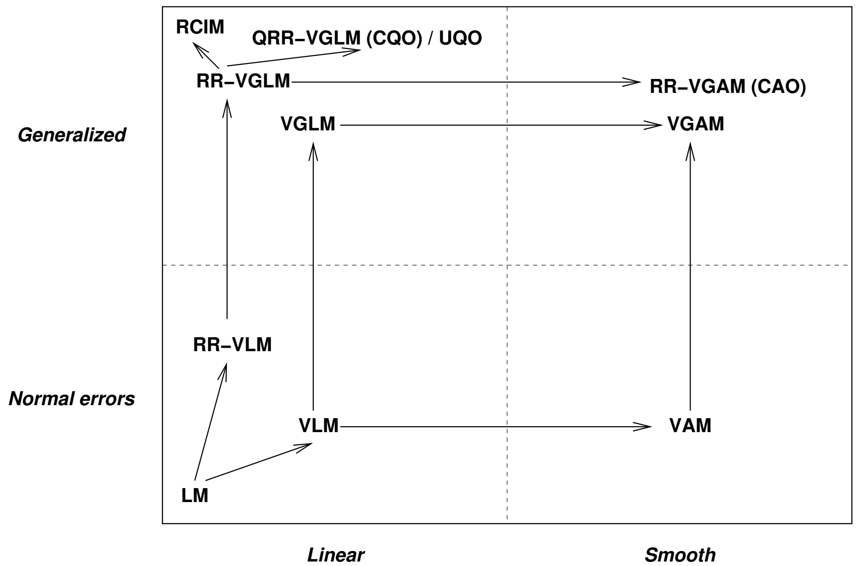


Fig. 1.2 Flowchart for different classes of models. Legend: LM = linear model, V = vector, G = generalized, A = additive, O = ordination, Q = quadratic, U = unconstrained, RCIM = row–column interaction model. See also Table 1.1. Apart from the LM, the models of the bottom half are more to be viewed as computational building blocks.

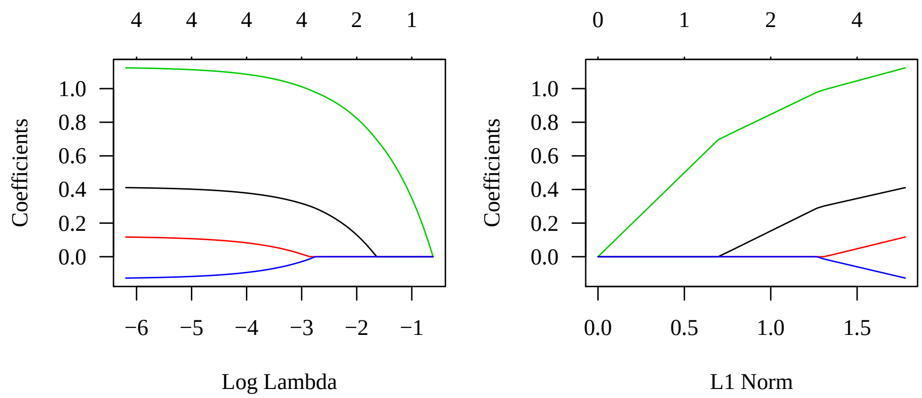


Fig. 1.3 Paths of the estimated LASSO coefficients in a LM fit. The response `log(los)` is regressed against the variables `admit` (black), `age75` (red), `procedure` (green) and `sex` (blue) and intercept, in the data frame `azpro` from COUNT. The first plot has $\log \lambda$ as its x-axis, whereas second plot has the quantity $\sum_{k=2}^{p} |\beta_k|$ in (1.43). The upper numbers are the number of variables in the model. Package `glmnet` is used here.

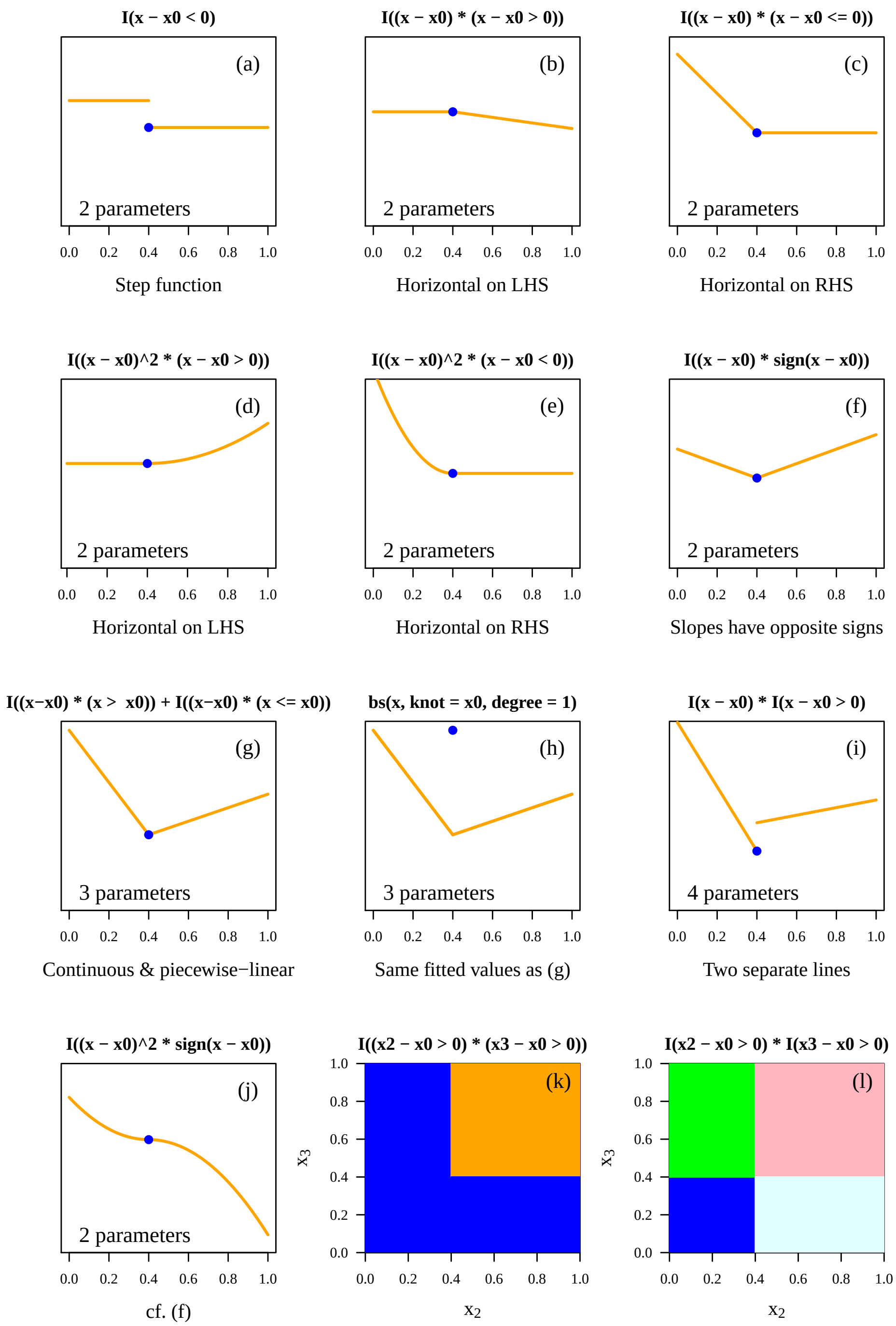


Fig. 1.4 Some functional forms derived from the S formula language based on elementary functions and operators. The formula heads each plot, followed by footnotes. The number of parameters in the regression is given. An intercept term is assumed in all—and the blue ● point indicates its value at the location $x = x_0$. Plots (k)–(l) are contour images with various colours denoting different fitted values (blue for the intercept). The value `x0` $= x_0 = 0.4$ here, and variables x, x_2 and x_3 are defined on the unit interval. Function `bs()` resides in the `splines` package.

Chapter 2
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

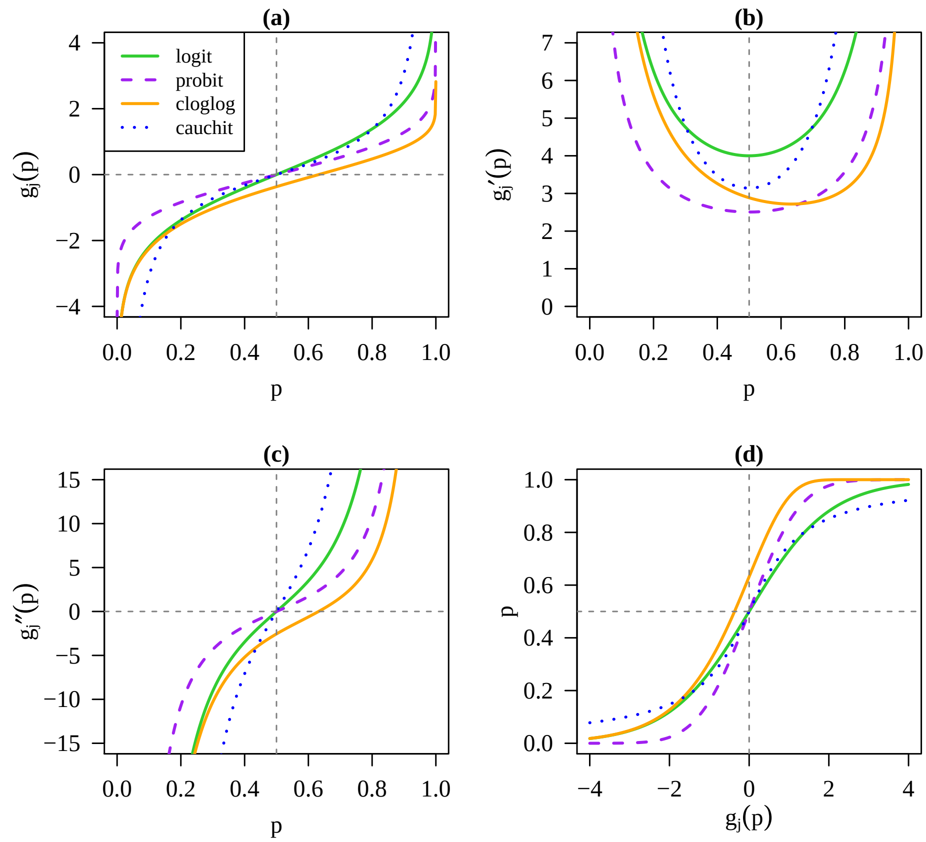


Fig. 2.1 Properties of some common link functions g_j suitable for a probability. (a) $g_j(p)$; (b) $g_j'(p)$; (c) $g_j''(p)$; (d) $g_j^{-1}(p)$. The legend in (a) is common for all plots. The calls to (a)–(c) are of the form `link.function(p, deriv = d)` for `d = 0`, `1` and `2` (Table 1.2).

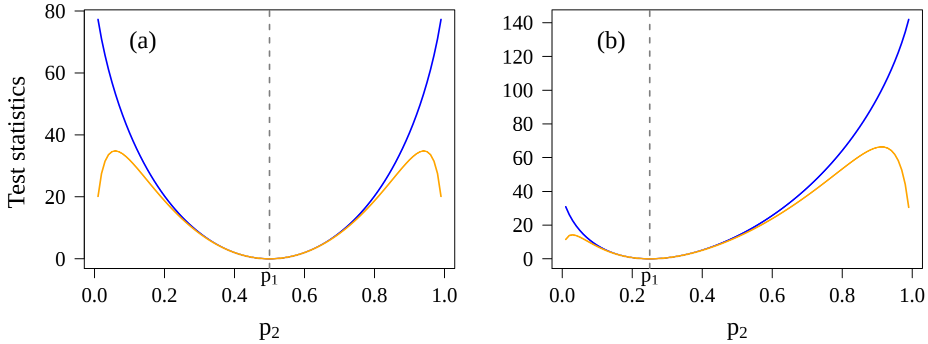


Fig. 2.2 Wald (orange) and likelihood ratio test (blue) statistics plotted against p_2, for: (a) $p_1 = 0.5$, (b) $p_1 = 0.25$ (vertical dashed lines). Actually, the Wald statistic here is the square of the usual Wald statistic, and $p_2 = 0.01(0.01)0.99$ is discrete. The data follows Hauck and Donner (1977).

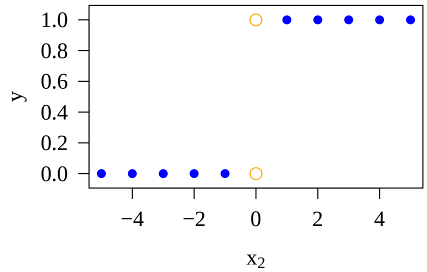


Fig. 2.3 Completely separable data (blue circles). Adding the two orange hollow points results in quasi-completely separable data. The logistic regression estimate of the slope will tend to infinity in both cases.

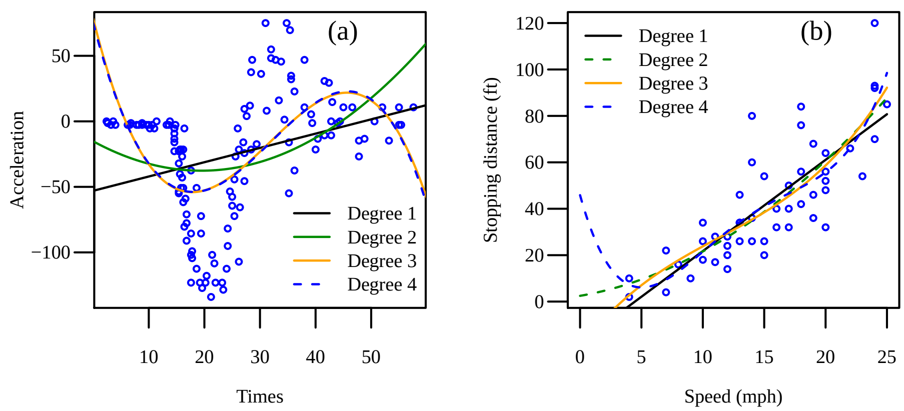


Fig. 2.4 Polynomials of degree 1–4 fitted to two data sets. (a) `mcycles` from MASS. (b) `cars` from datasets.

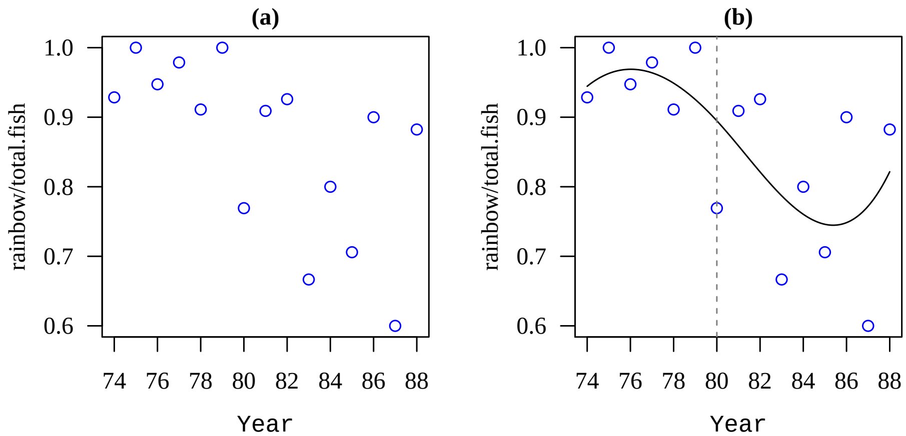


Fig. 2.5 (a) Proportion of fish caught that are rainbow trout from Lake Otamangakau (`lakeO`) caught by an angler who frequented the spot. The variable `Year` is `year-1900`. (b) Smoothing the same data with a cubic regression spline (truncated power basis) with one knot located at the year 1980. A boundary effect on the RHS is evident.

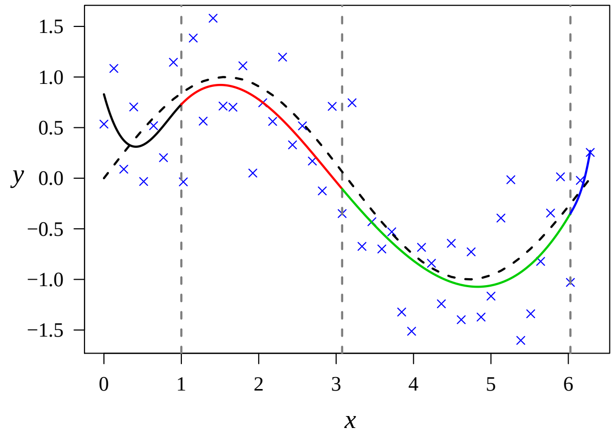


Fig. 2.6 Smoothing some data with a regression spline (B-spline). Each segment of the spline is coloured differently. The term is effectively `bs(x, knots = c(1, 3.08, 6.03))`. The true function is the sine function (dashed) and $n = 50$.

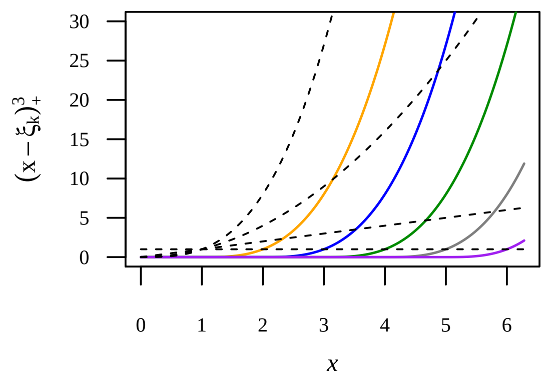


Fig. 2.7 Truncated power series basis for cubic splines (2.39) The black dashed lines are $1, x, x^2, x^3$. The coloured solid lines are $(x - \xi_k)^3_+$ for knots $\xi_1 = 1, \xi_2 = 2, \xi_3 = 3, \xi_4 = 4$ and $\xi_5 = 5$.

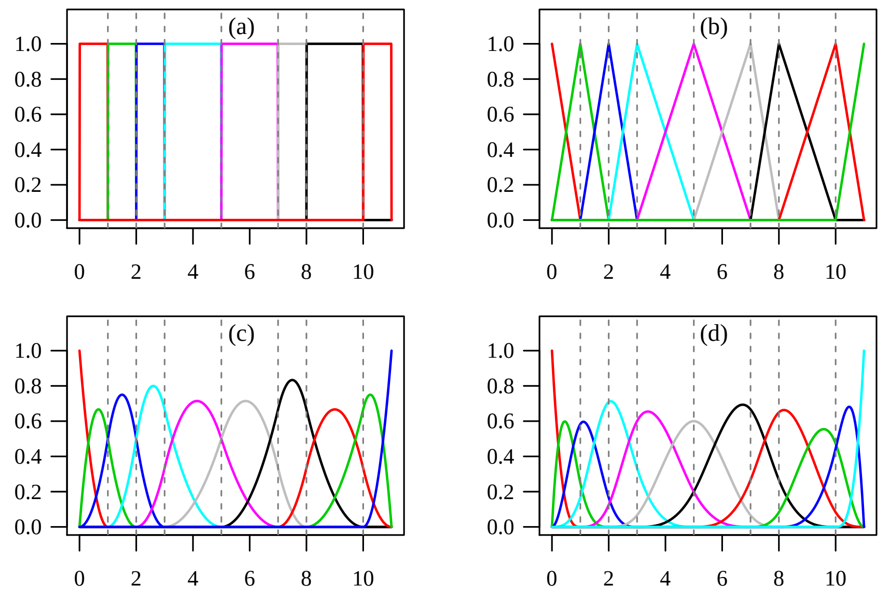


Fig. 2.8 B-splines of order 1–4 ((a)–(d)), where the interior knots are denoted by vertical lines. The boundary knots are at 0 and 11. The basis functions have been plotted from left to right.

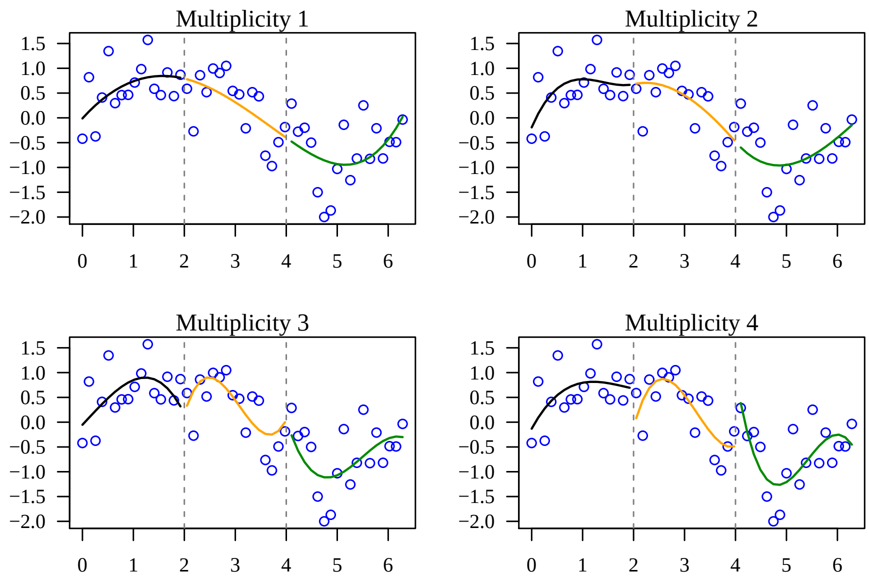


Fig. 2.9 Smoothing some data with cubic regression splines with knots of varying multiplicities. The knots are at $x = 2$ and 4.

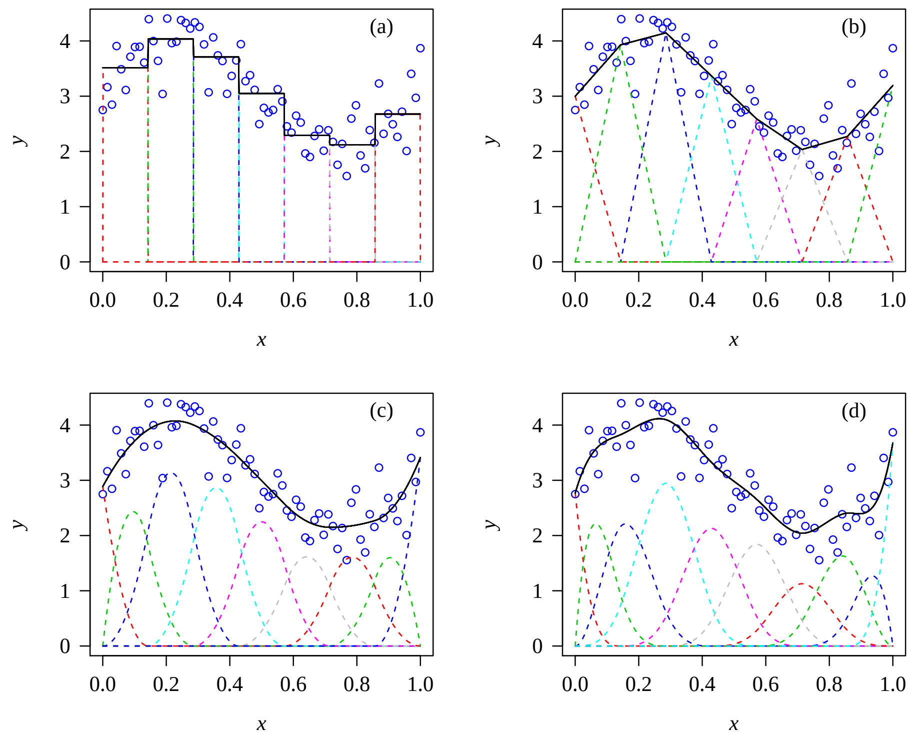


Fig. 2.10 (a)–(d) Linear combinations of B-splines of degrees 0–3 fitted to some scatter plot data; the formula is similar to (2.55) The knots are equally-spaced on the unit interval.

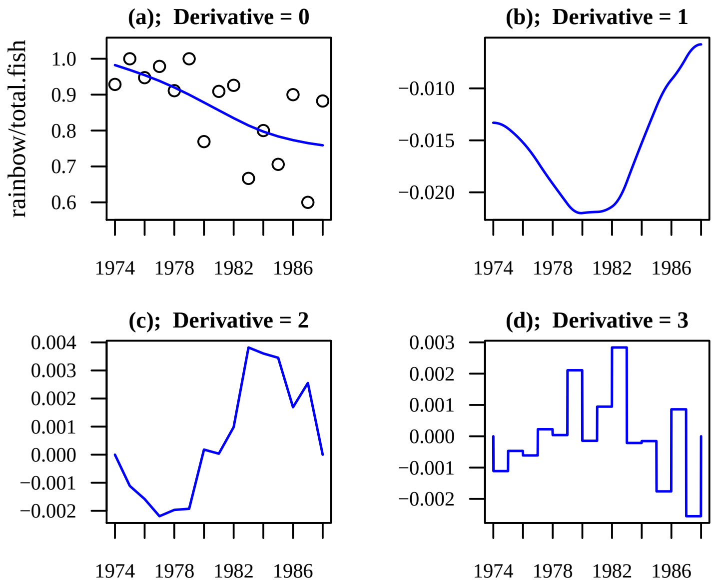


Fig. 2.11 (a) Cubic smoothing spline fitted to the proportion of fish caught that are rainbow trout from `lake0`. The x-axis is `year`. The smoother has 1 nonlinear degrees of freedom. (b)–(d) Derivatives of the smooth of orders 1–3. In contrast, Fig. 2.15 fits a local linear regression to these data.

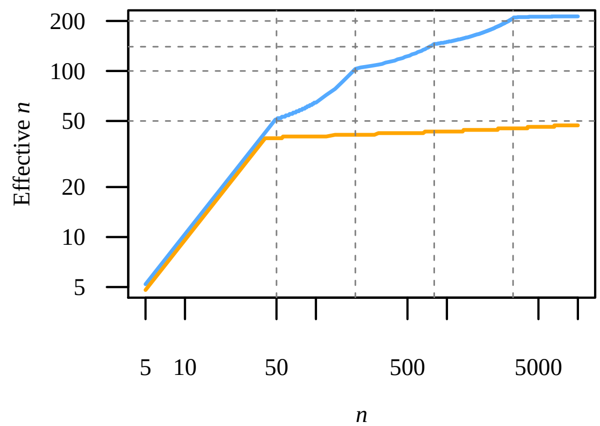


Fig. 2.12 O-splines: number of knots K selected from n unique x_i, for `smooth.spline()` is the top function. The lower function is (2.56) for `vsmooth.spline()`. Both axes are on a logarithmic scale. The top function intersects with the dashed lines at $(50, 50)$, $(200, 100)$, $(800, 140)$ and $(3200, 200)$; logarithmic interpolation is used for other n values.

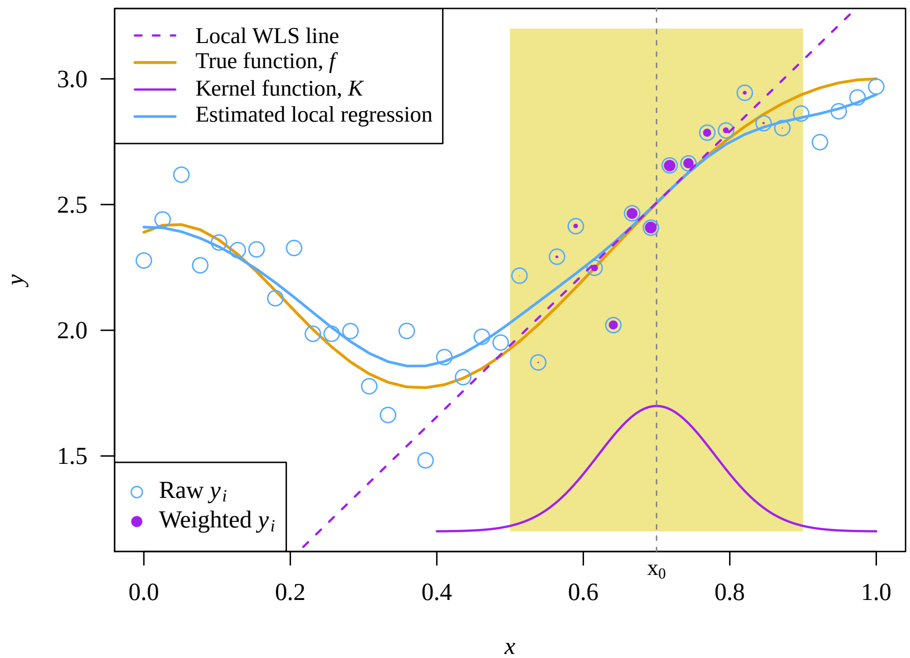


Fig. 2.13 Local linear regression with $n = 40$ points. The kernel weights have been divided by 10 for scaling purposes. The vertical line is at the target point $x_0 = 0.7$ so that three curves/lines intersect at $(x_0, \widehat{f}(x_0))$. The shaded region is effectively the window for computing $\widehat{f}(x_0)$.

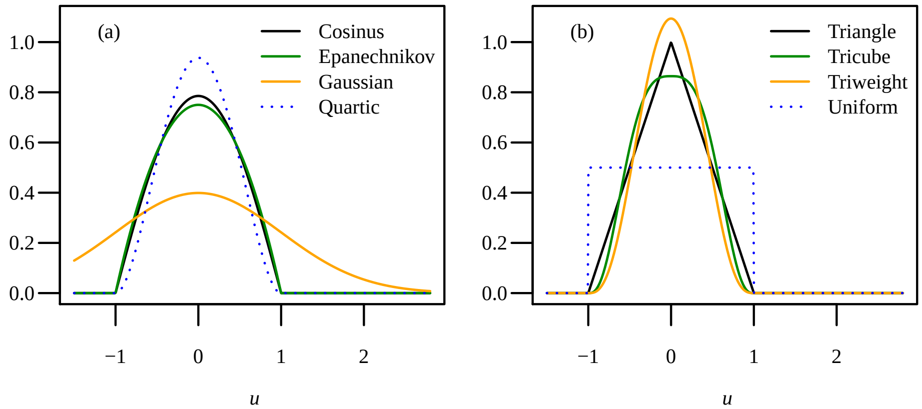


Fig. 2.14 Kernel functions from Table 2.2. All but one has compact support.

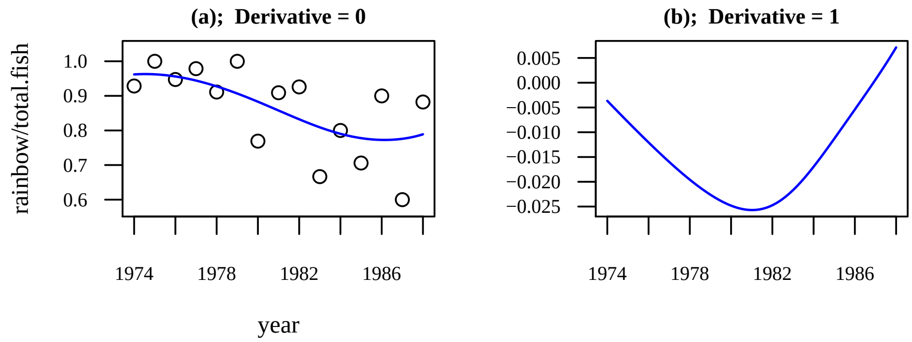


Fig. 2.15 (a) Local linear regression fitted to the proportion of fish caught that are rainbow trout from `lake0`. The smoother has a bandwidth of $h = 2.5$ and uses the Gaussian kernel function. (b) $\widehat{f}'(x)$. In contrast, Fig. 2.11 fits a cubic smoothing spline to these data.

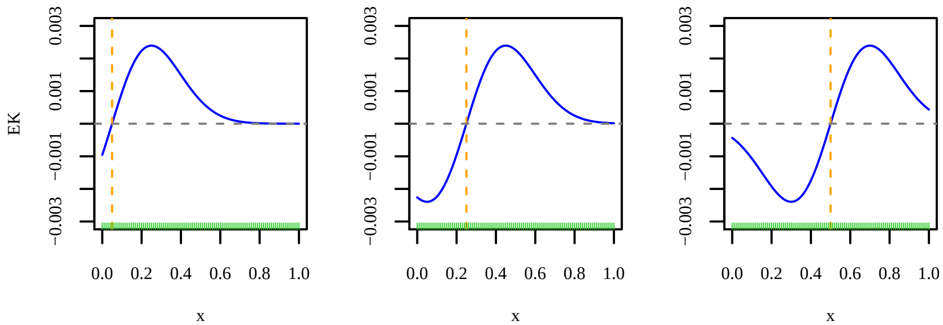


Fig. 2.16 Asymptotic equivalent kernel for $\widehat{f}'(x)$ from (2.82) The 101 x_i are equally-spaced on $[0, 1]$ (as shown by the rugplot). The x_0 values are 0.05, 0.25, 0.5 (vertical dashed lines), and the bandwidth is 0.2. The kernel function $K = \phi(\cdot)$.

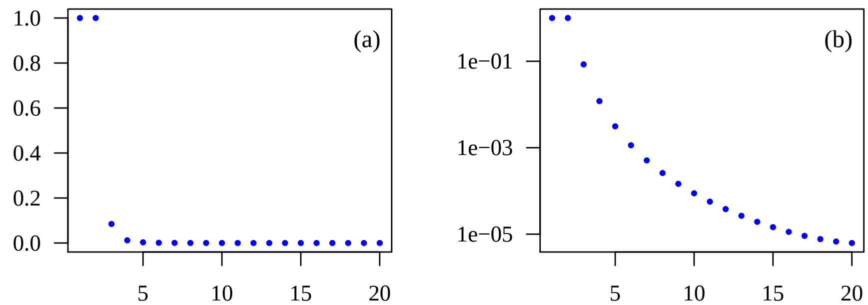


Fig. 2.17 (a) Eigenvalues of the smoother matrix of a cubic smoothing spline. Here, $n = 20$ and the x_i are equidistant on $[0, 1]$. (b) The same on a logarithmic scale. The two unit eigenvalues correspond to constant and linear functions. The corresponding eigenvectors are plotted in Fig. 2.18.

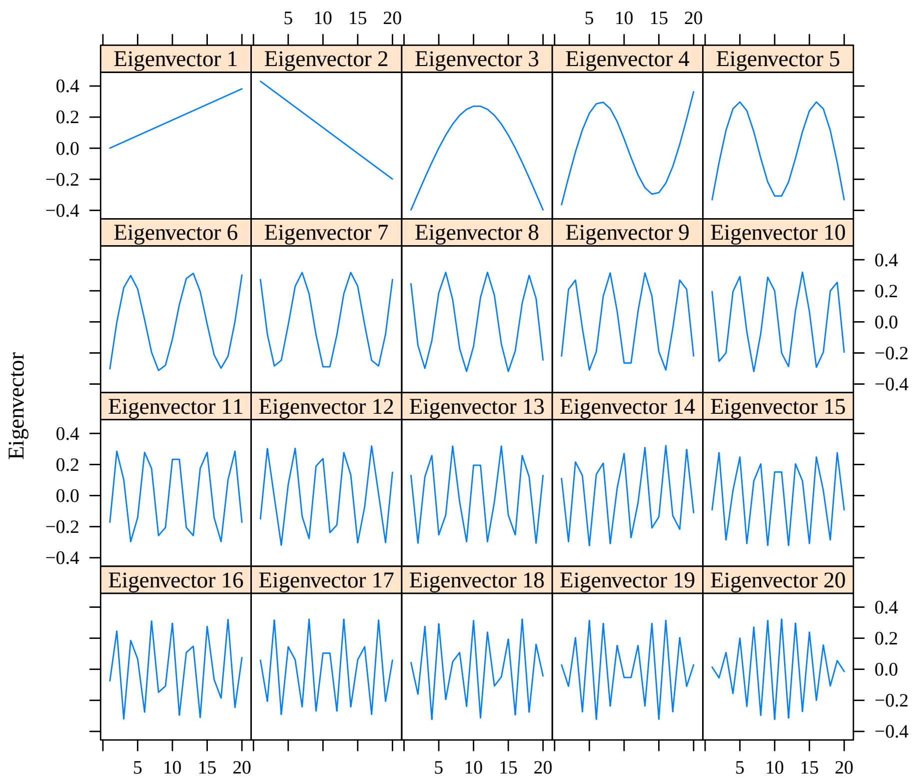


Fig. 2.18 Successive eigenvectors corresponding to the eigenvalues of Fig. 2.17.

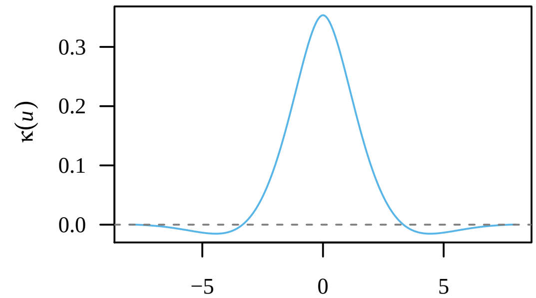


Fig. 2.19 Equivalent kernel of a cubic spline, $\kappa(u)$ (Eq. (2.92))

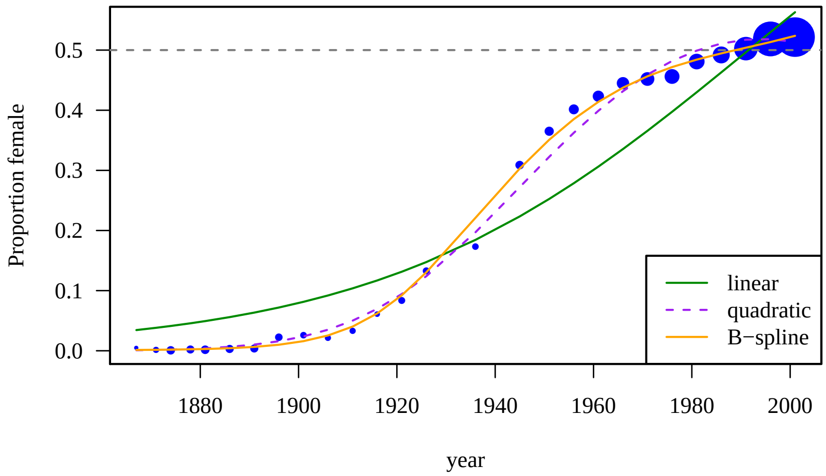


Fig. 2.20 Fitted values from some logistic regression models applied to `chinese.nz`. The response is the proportion of New Zealand Chinese who are female. The terms are `year`, `poly(year, 2)`, `bs(year, 4)`. Area sizes of the points are proportional to the number of people.

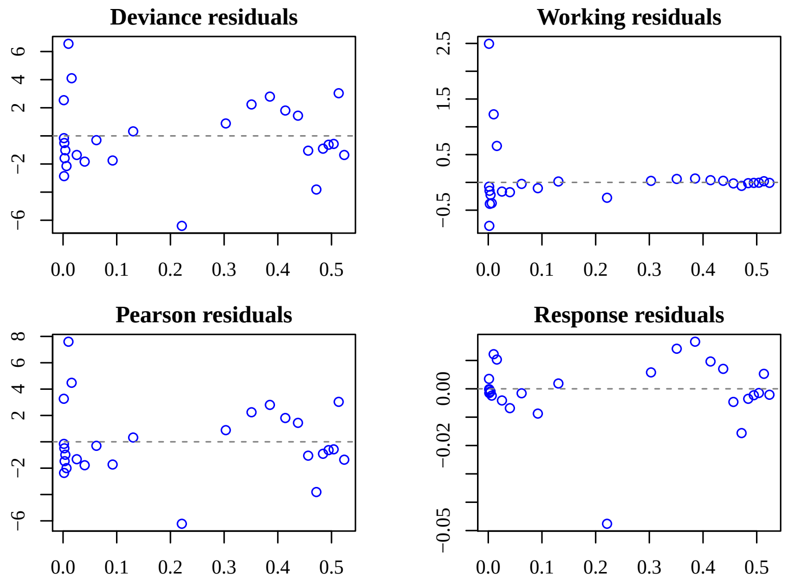


Fig. 2.21 Four residual types for the regression spline fit of Fig. 2.20. The fitted values are plotted on the x-axis.

Chapter 3
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

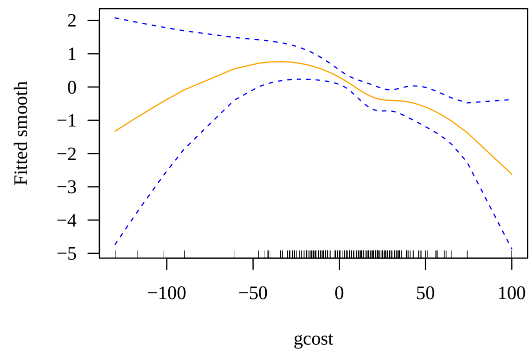


Fig. 3.1 An extension of the fitted model `fit.travel` from Sect. 14.2.1, where the cost variable is smoothed with regression splines. The variable `gcost` stands for 'generalized cost'.

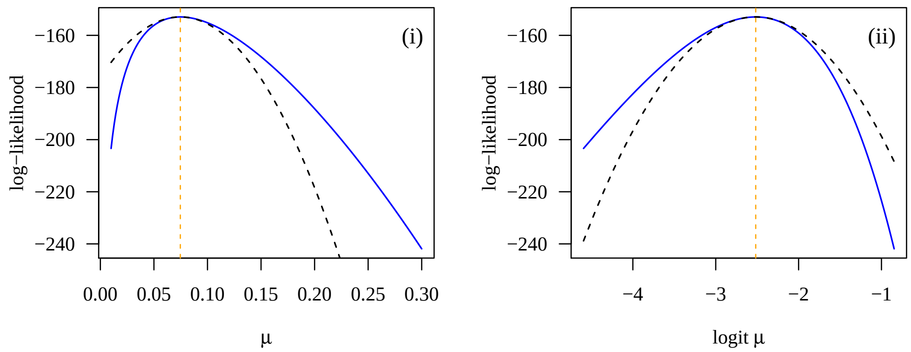


Fig. 3.2 Log-likelihood ℓ as a function of (i) μ and (ii) $\eta = \text{logit}\,\mu$, for the `V1` data frame. The model is a logistic regression with 3 or more hits defining 'success'. The vertical dashed lines are at the MLE $\widehat{\mu}$. The dashed curves are the quadratic approximation to ℓ at $\widehat{\mu}$.

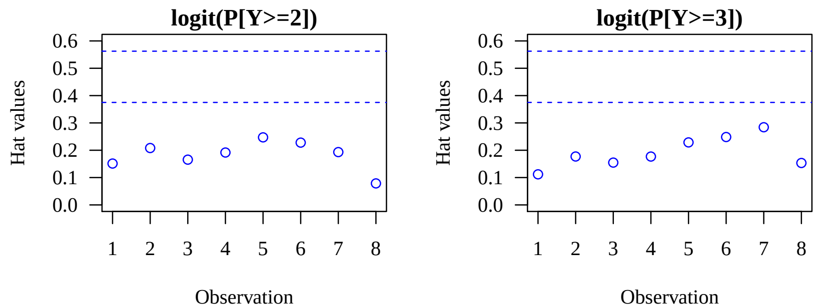


Fig. 3.3 Hat values from a proportional odds model fitted to the severity of pneumoconiosis data set in coalminers, called `pneumo`.

Chapter 4

Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

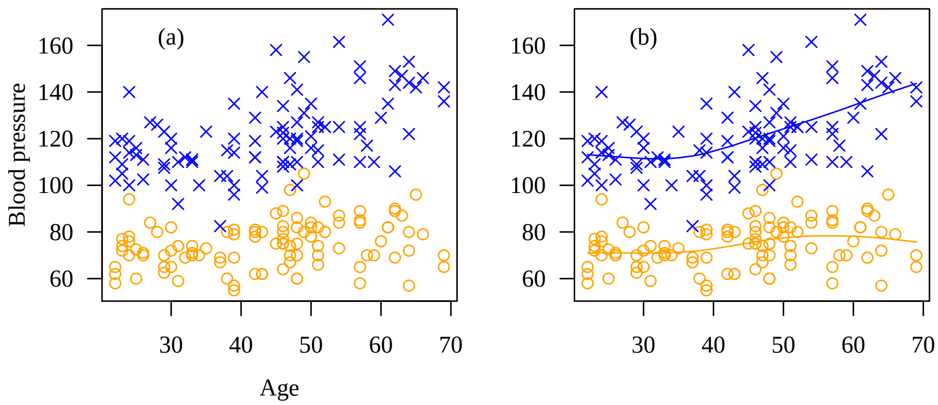


Fig. 4.1 (a) Scatter plot of diastolic (○) and systolic (×) blood pressures (mm Hg) versus age, for a random sample of 100 European-type females from `xs.nz`. (b) A vector smoothing spline fit overlaid on the same. Each component function has 2 effective nonlinear degrees of freedom (ENDF).

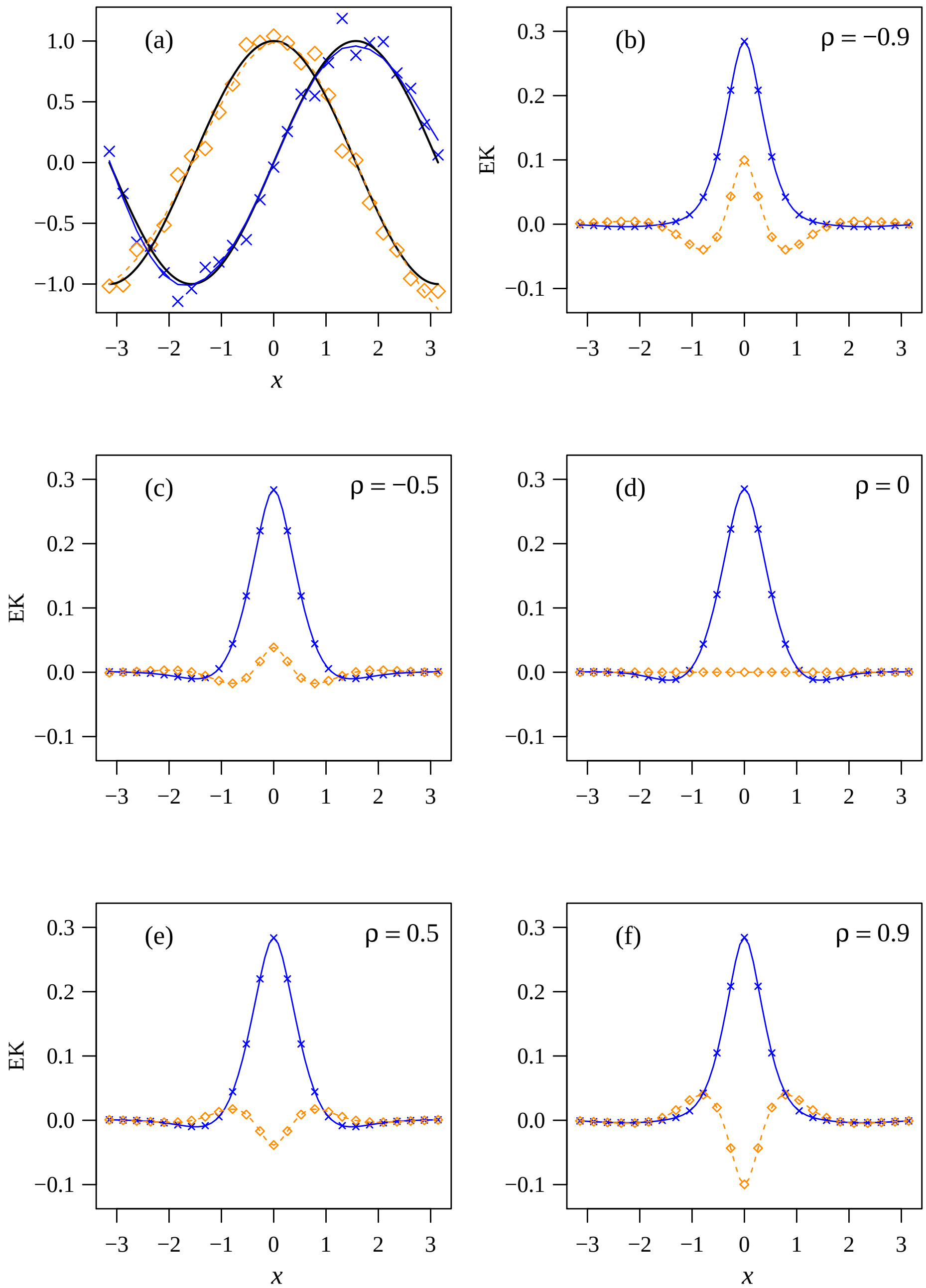


Fig. 4.2 (a) Data used from (4.6). The black curves are the true functions; (b)–(f) Equivalent kernels based on (4.6) for various values of ρ. The row of the influence matrix corresponds to $\widehat{f}_1(x = 0)$, and is midway between the boundaries.

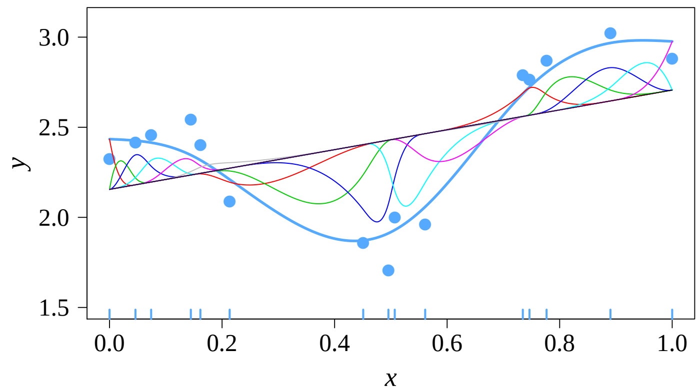


Fig. 4.3 O-smoothing spline (blue curve) fitted to scatter plot data ●, which is equal to a least squares fit (black line) plus the sum of individual B-spline basis functions; cf. Fig. 2.8. It corresponds to the decomposition (4.12). The rugplot denotes the position of the x_i.

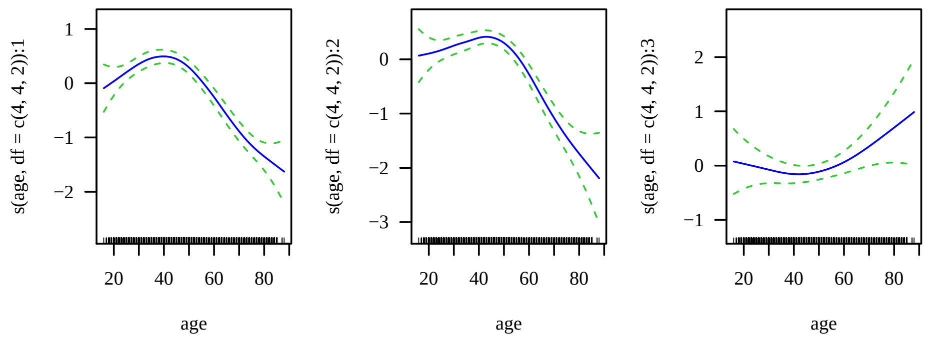


Fig. 4.4 A bivariate odds ratio model fitted as a VGAM to a subset of female Europeans from `xs.nz`, with household cat and dog ownership as the responses. The plots are the $\widehat{f}_{(j)2}(x_2)$.

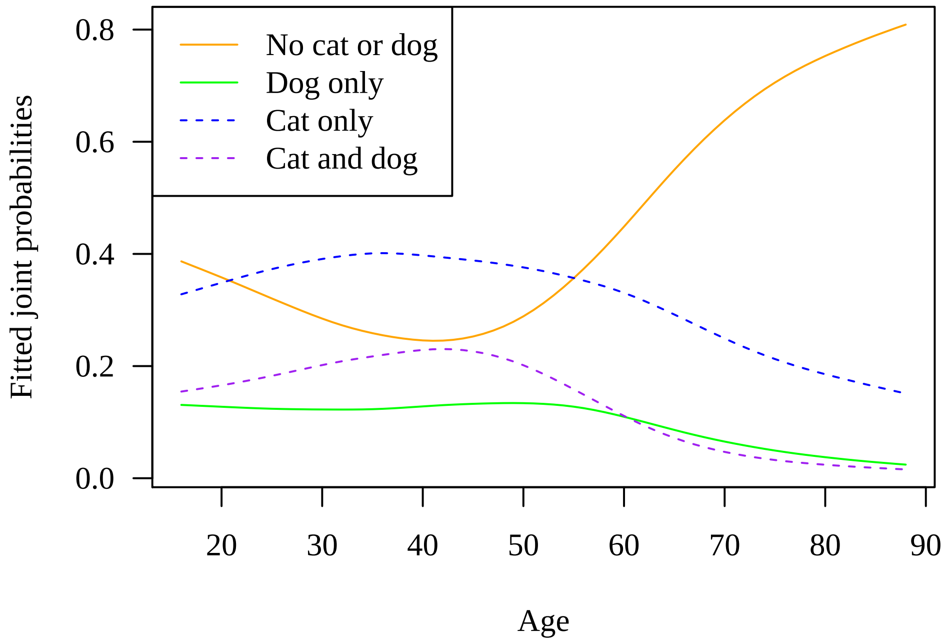


Fig. 4.5 Joint probabilities from `fit.cd1`.

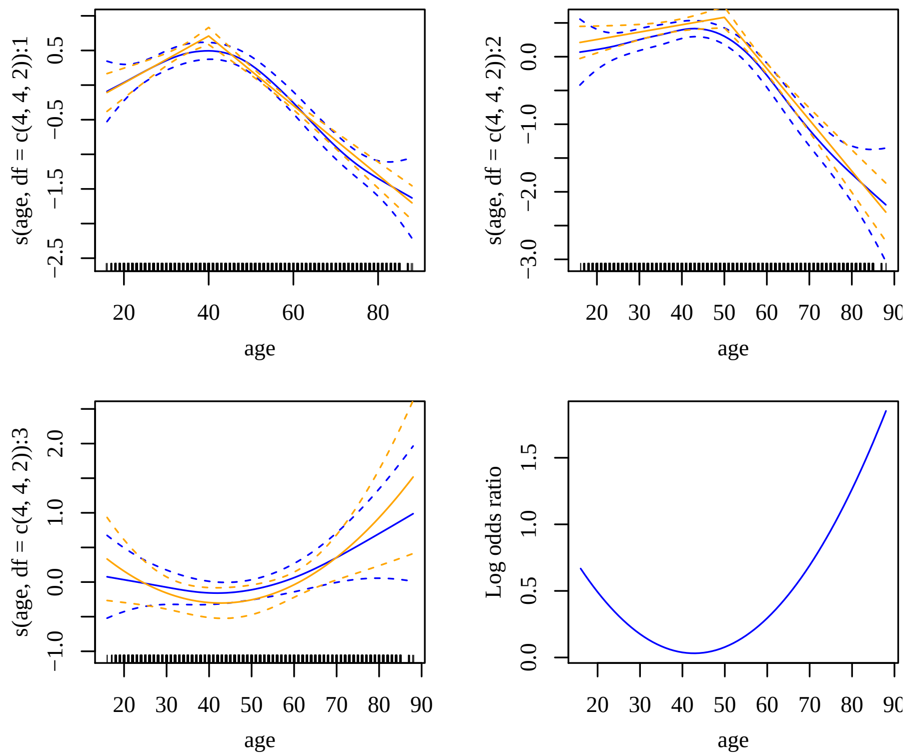


Fig. 4.6 The first three plots are the fitted component functions overlaid; the models are `fit1.cd` are `fit3.cd`. The fourth plot is the estimated log odds ratios of `fit3.cd`.

Chapter 5

Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

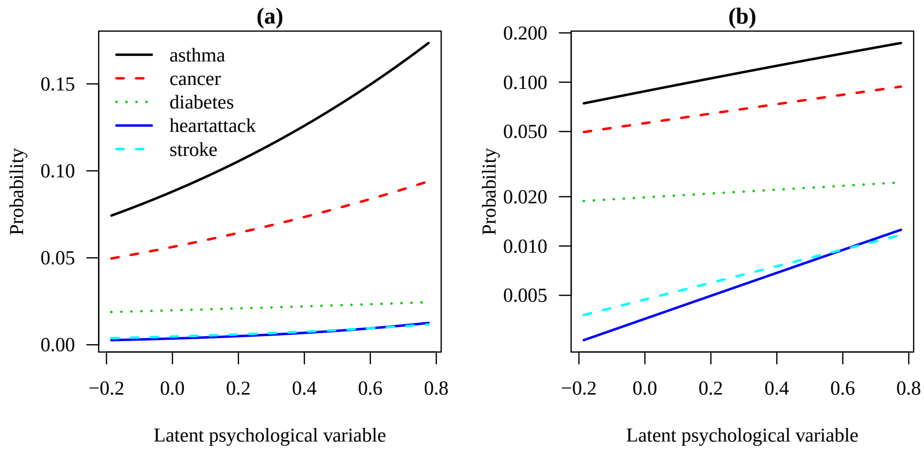


Fig. 5.1 (a) RR-VGLM-binomial model applied to several disease responses. The x-axis is $\widehat{\nu}$, a linear combination of 11 binary psychological variables. The y-axis is disease prevalence. (b) The same on a log scale.

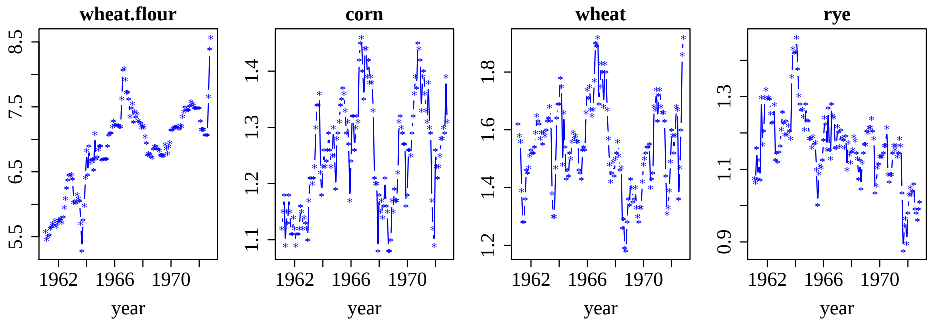


Fig. 5.2 Monthly average prices of Grain series, January 1961–October 1972, in data frame `grain.us`.

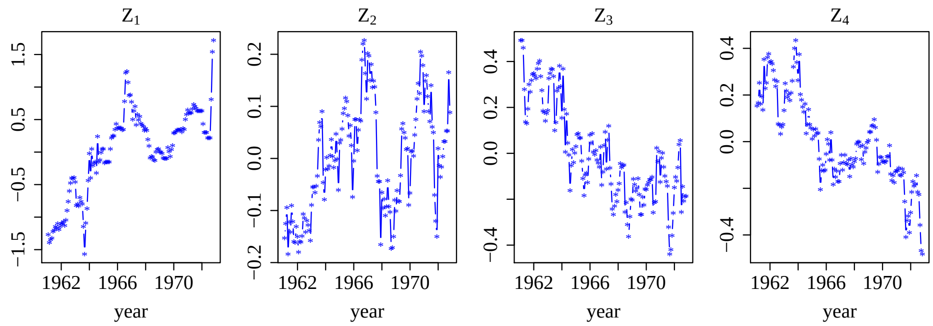


Fig. 5.3 Canonical variables of the Grain Price series, January 1961–October 1972.

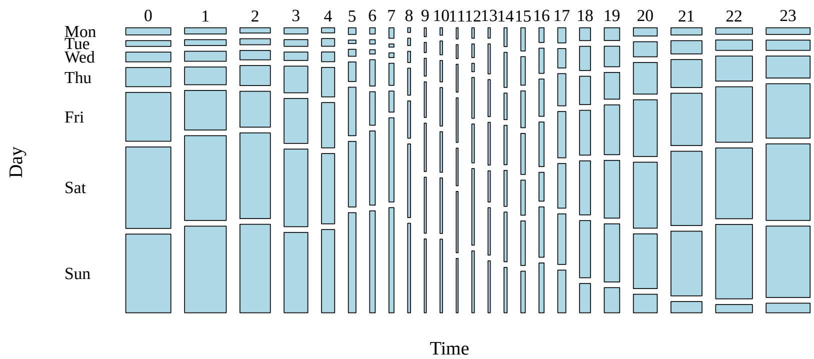


Fig. 5.4 Mosaic plot of `alcoff`; the area sizes are proportional to the counts.

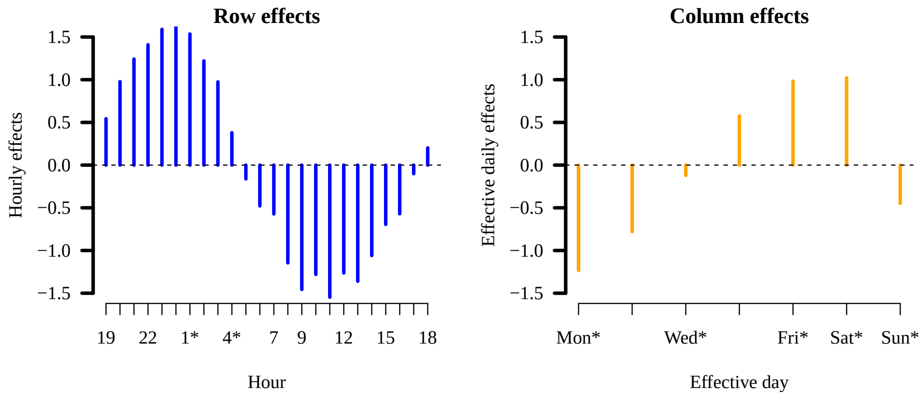


Fig. 5.5 Hourly and effective daily effects of a rank-0 Goodman's RC model fitted to `alcoff`. This is output from `plot(grc0.alcoff)`.

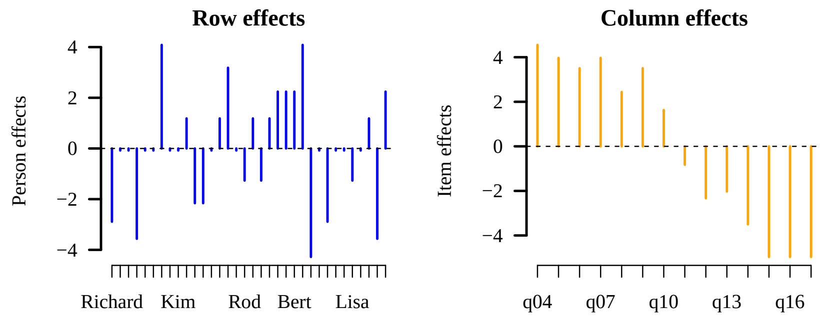


Fig. 5.6 Rasch fixed effects model to `exam1`, Eq. (5.29). Only a few people and items are labelled.

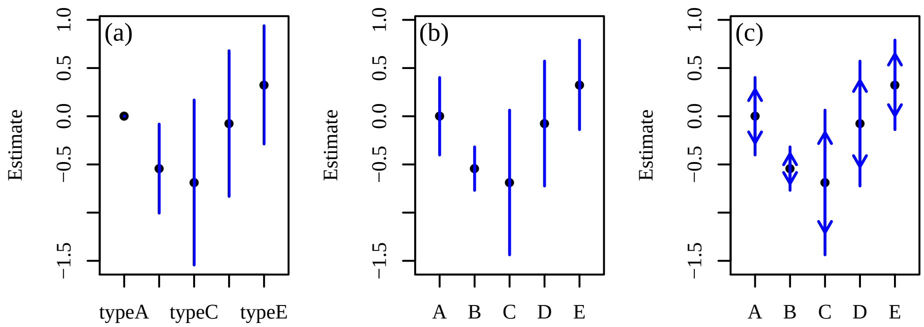


Fig. 5.7 Quasi-variances computed for the `ships` data in MASS. The fitted model is a quasi-Poisson-GLM fitted to the ship data (McCullagh & Nelder, 1989) with respect to ship types (A–E) on the damage rate on a log scale. (a) Confidence intervals for contrasts with type A ships based on conventional standard errors; (b) Comparison intervals based on quasi-variances; (c) 5% LSD intervals (arrows) based on quasi-variances overlaid on (b). For (a)–(c), the formulas are $\widehat{\beta}_i \pm 2\,\mathrm{SE}(\widehat{\beta}_i)$, $\widehat{\beta}_i \pm 2\sqrt{q_i}$ and $\widehat{\beta}_i \pm z(0.025)\sqrt{q_i/2}$, respectively.

Chapter 6
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

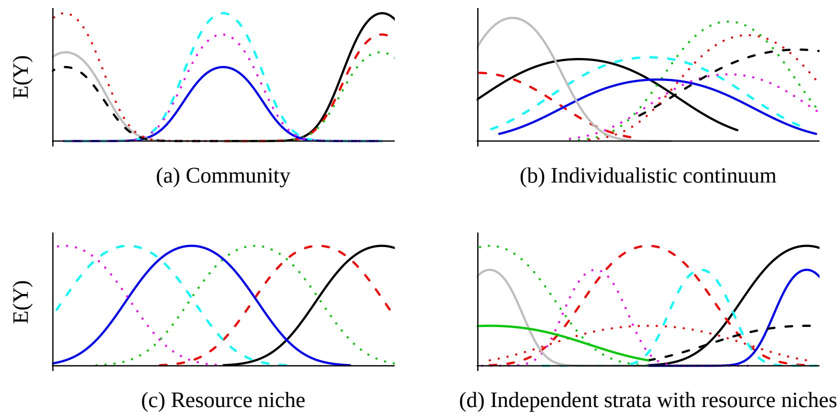


Fig. 6.1 Four hypothetical models for community ecology, along a gradient or latent variable. Plot (c) corresponds to a species packing model.

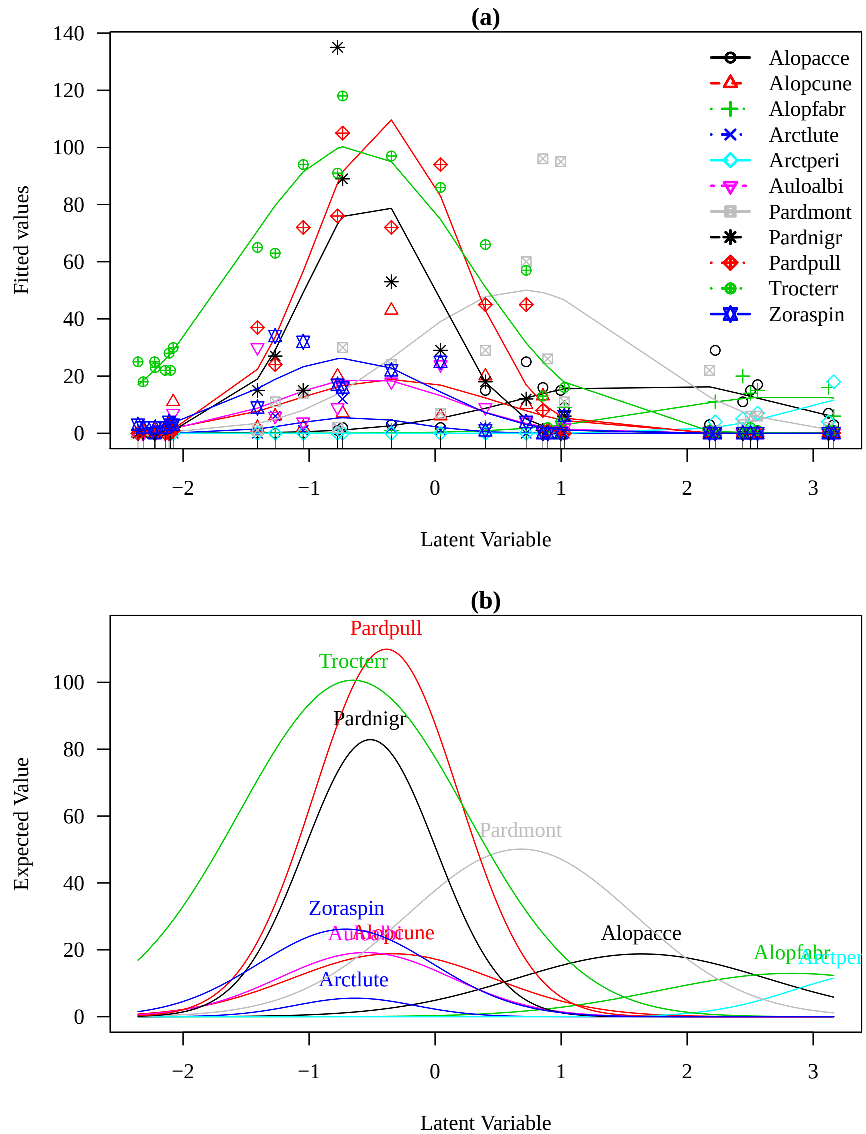
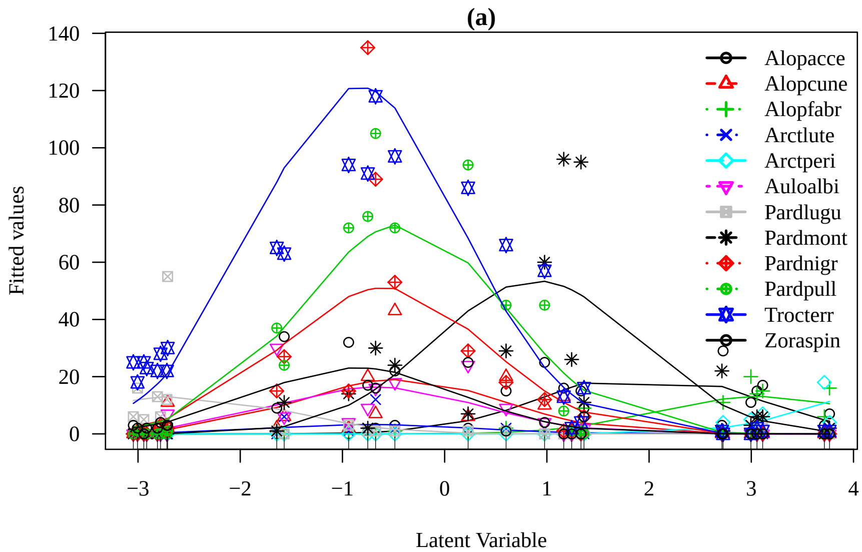


Fig. 6.2 Ordination of 11 species from the `hspider` data frame; a Poisson model with unequal tolerances, called `p1ut.hs.2`. (a) `lvplot()` output. (b) `persp()` output, which is a perspective plot and a 'continuous' version of (a).



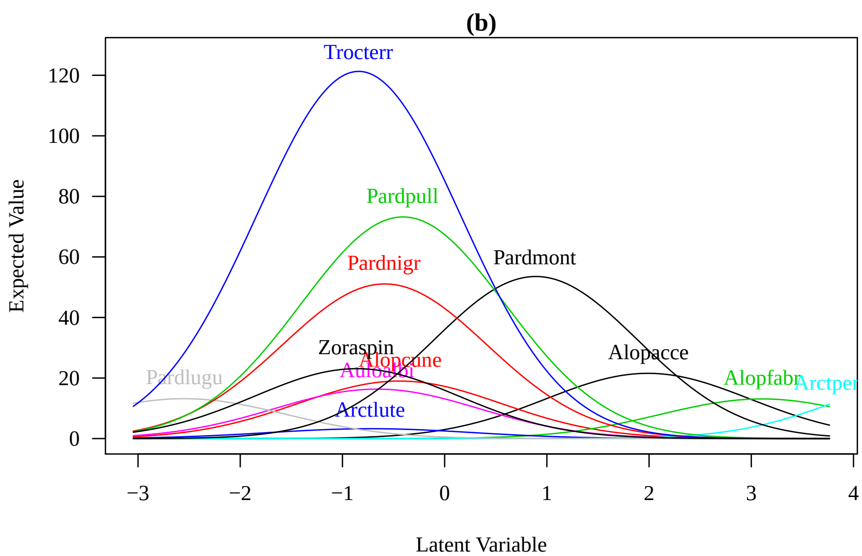


Fig. 6.3 Ordination of 12 species from the `hspider` data frame; a Poisson model with equal tolerances, called `p1et.hs`. (a) `lvplot()` output. (b) `persp()` output, which is a perspective plot and a 'continuous' version of (a).

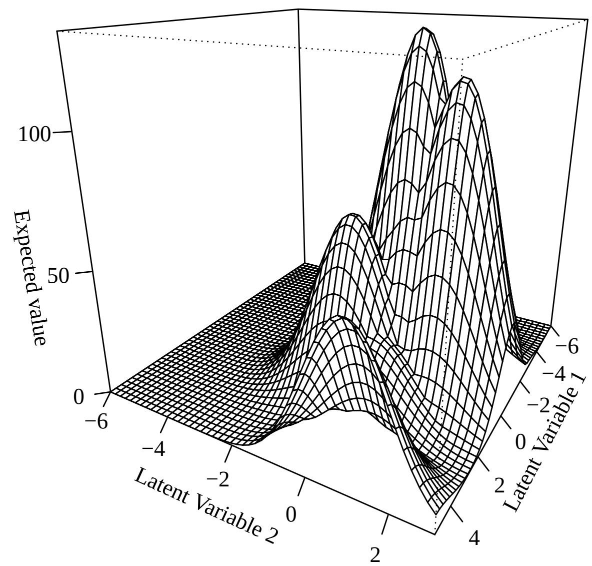


Fig. 6.4 Perspective plot of a rank-2 CQO fitted to the `hspider` data frame, called `p2et.hs`. All 12 species are fitted, using an equal-tolerances Poisson model, but only the response surfaces of the dominant species may be seen. See also Fig. 6.6 for an ordination diagram of the above.

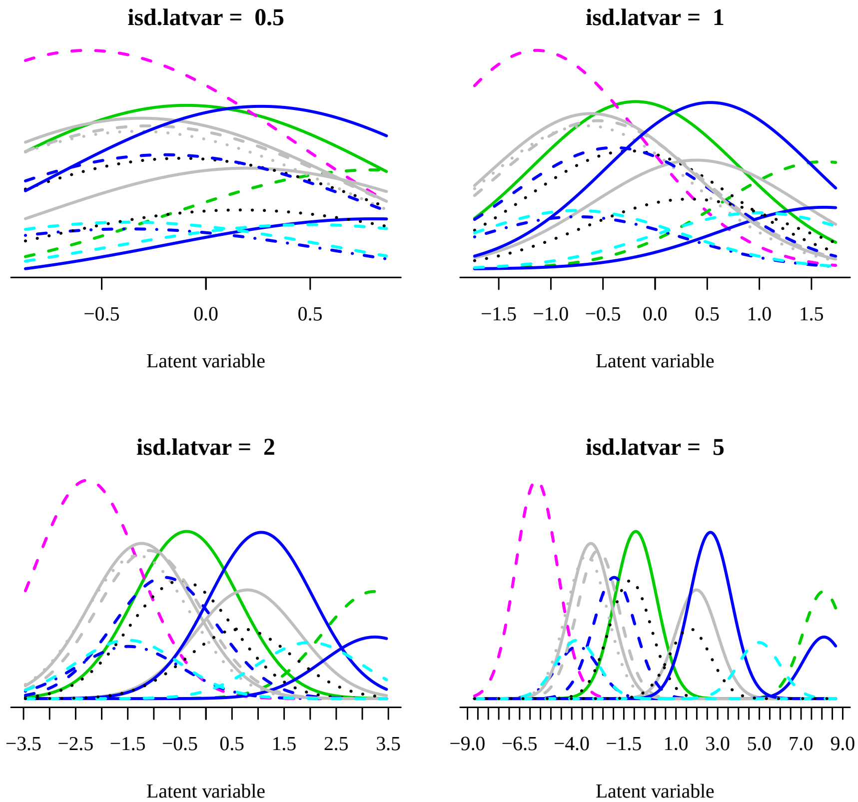


Fig. 6.5 The effect of the argument `isd.latvar` on the site scores. All response curves have unit tolerances (because $\mathbf{T}_s = \mathbf{I}_R \ \forall s$), and the optimums are located the same relative distance from each other. The site scores are uniformly distributed over the latent variable space, and have been scaled to have a standard deviation `isd.latvar`. The tick marks are at the same values.

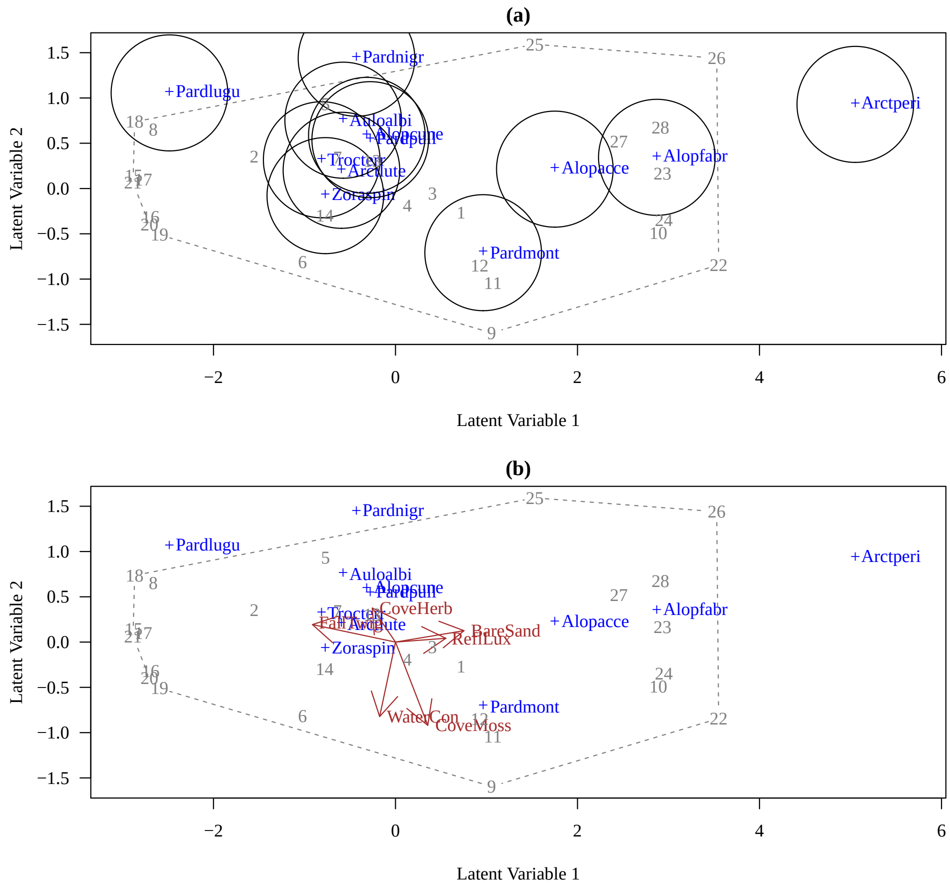


Fig. 6.6 Ordination diagrams of 12 species from the `hspider` data frame; the Poisson model `p2et.hs` with equal tolerances. A convex hull surrounds the site scores. In (a), the circles indicate the abundance of each species at 95% of its maximum abundance. In (b), the arrows display the contribution of each environmental variable towards each of the ordination axes.

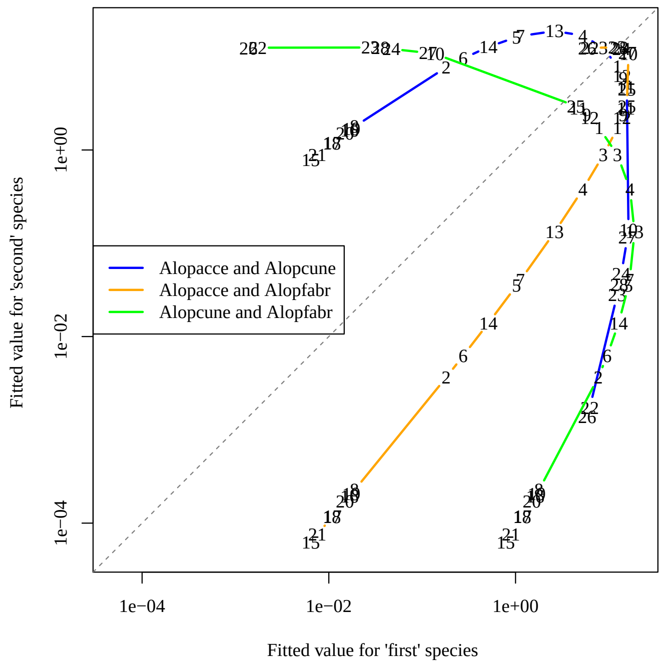


Fig. 6.7 Trajectory plot of three hunting spiders species. A rank-1 Poisson CQO is fitted to these. Site numbers have been placed on each curve.

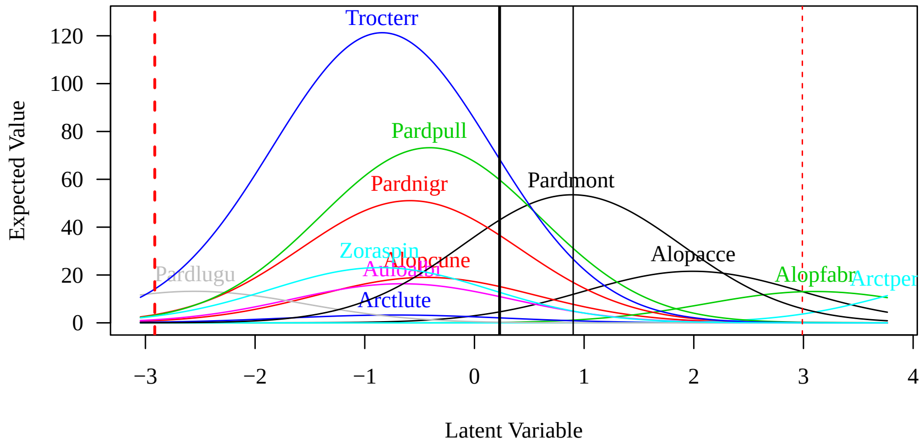


Fig. 6.8 Two calibrated sites from a rank-1 equal-tolerances Poisson QRR-VGLM fitted to the hunting spiders data without those 2 sites. The thick vertical lines are the CQO sites scores $\widehat{\nu}_i$, and the thinner vertical lines are the calibrated sites scores $\widetilde{\nu}_i$. Each i has the same colour and line type.

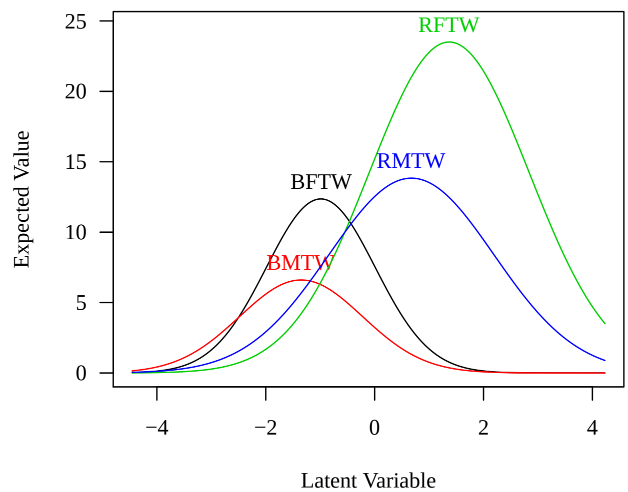


Fig. 6.9 Perspective plot of an unequal-tolerances Poisson CQO model fitted to the `trapO` trout data. Legend: "B" = brown trout, "R" = rainbow trout, "F" = female, "M" = male. The trap was located at the Te Whaiau Trap, hence the "TW". The `BFTW` response curve has unit tolerance.

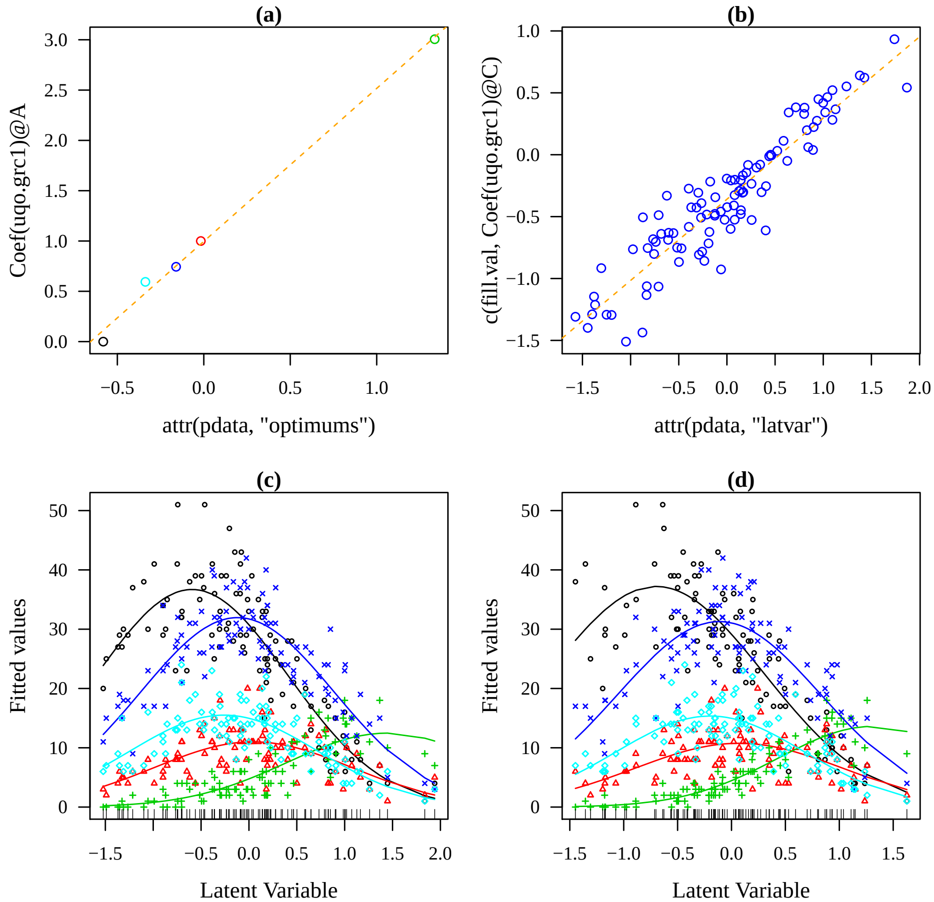


Fig. 6.10 (a)–(b) compares the UQO solution with the truth for a rank-1 Poisson simulated data set. (a) Estimated optimums $\widehat{u}_j$ versus u_j. (b) Estimated site scores $\widehat{\nu}_i$ versus ν_i. (c) CQO fitted to the original data. (d) CQO fitted to the scaled UQO site scores. In (a)–(b) the dashed orange line is a simple linear regression through the points.

Chapter 7

Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

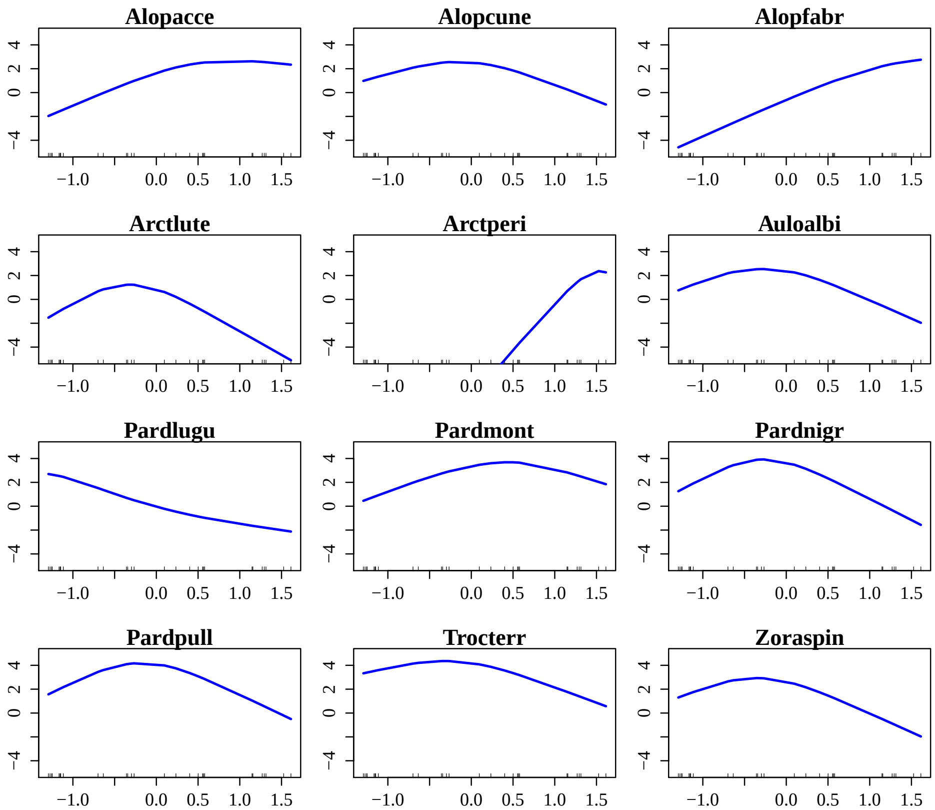


Fig. 7.1 Estimated (uncentred) component functions of `p1cao.hs`, a rank-1 Poisson RR-VGAM fitted to the `hspider` data. The x-axis is $\widehat{\nu}$.

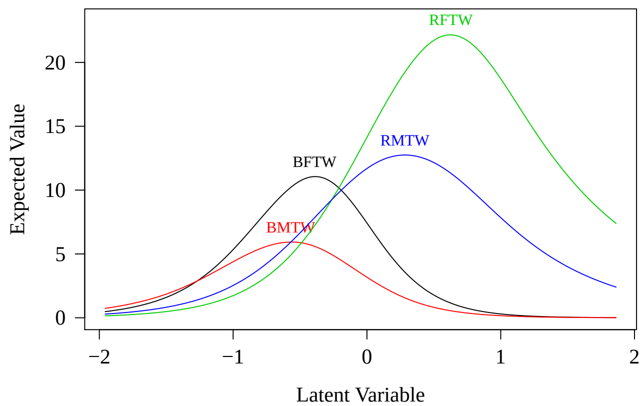


Fig. 7.2 Perspective plot of a rank-1 Poisson RR-VGAM fitted to the `trap0` data. The latent variable of this CAO, which is predominantly the day of the year, has unit variance. The responses are combinations of male and female rainbow and brown trout, all captured at the Te Whaiau Trap of Lake Otamangakau. See also Fig. 6.9.

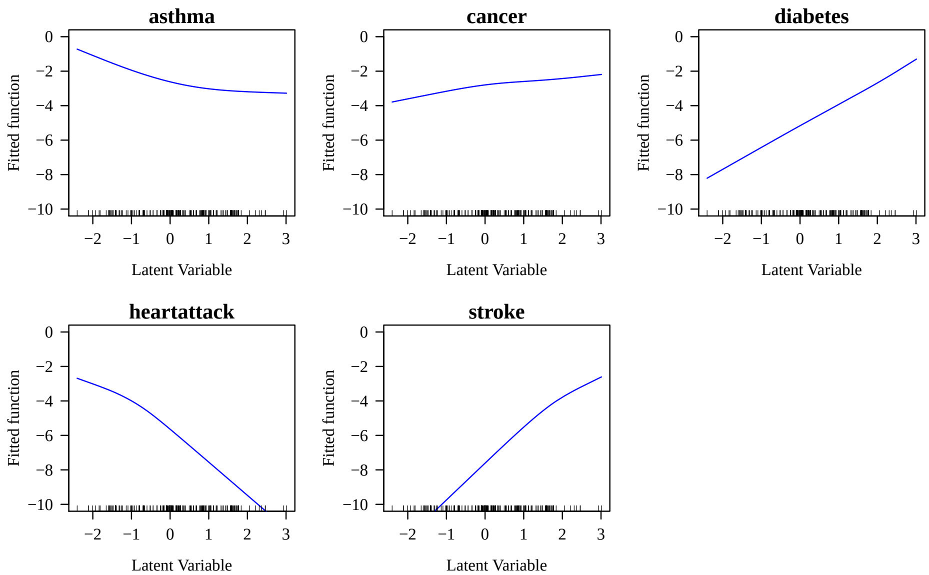


Fig. 7.3 Component functions (uncentred) of a rank-1 binomial RR-VGAM fitted to a subset of the `xs.nz` data. The latent variable of this CAO is a linear combination of 11 binary psychological variables.

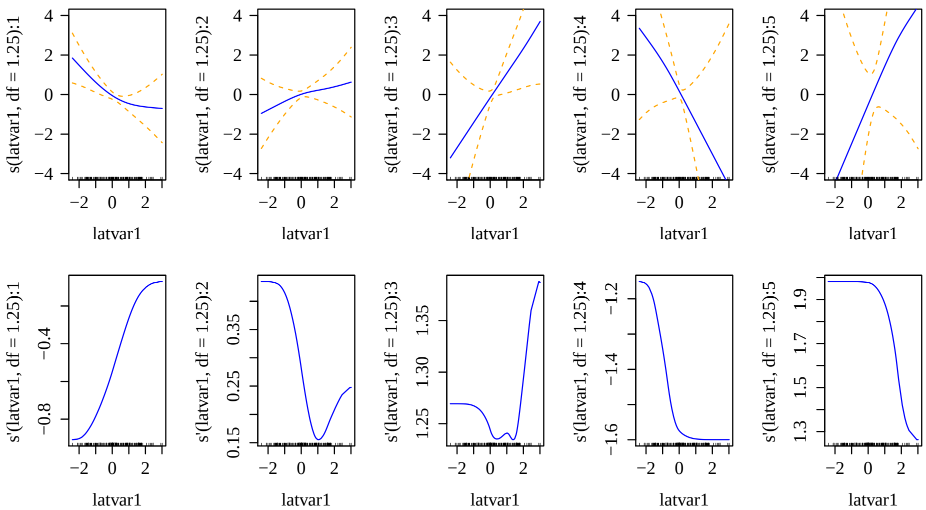


Fig. 7.4 Top row: component functions (centred) of a rank-1 binomial RR-VGAM fitted to a subset of the `xs.nz` data, cf. Fig. 7.3. The latent variable of this CAO is a linear combination of 11 binary psychological variables. Bottom row: first derivatives of the respective fitted functions.

Chapter 8
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

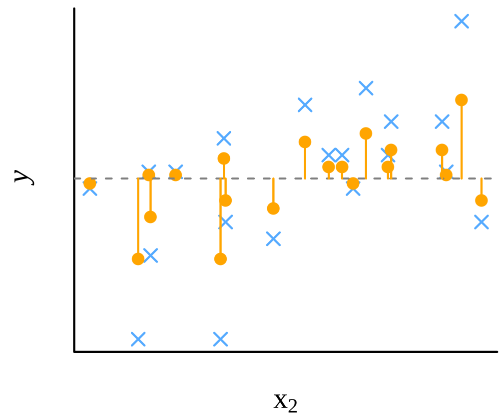


Fig. 8.1 Response y_i (×) in a simple linear regression, with shrinkage initial values (●) based on $s = \frac{1}{2}$ in (8.5). The dashed horizontal line is at $\overline{y}$, and s is argument `ishrinkage`.

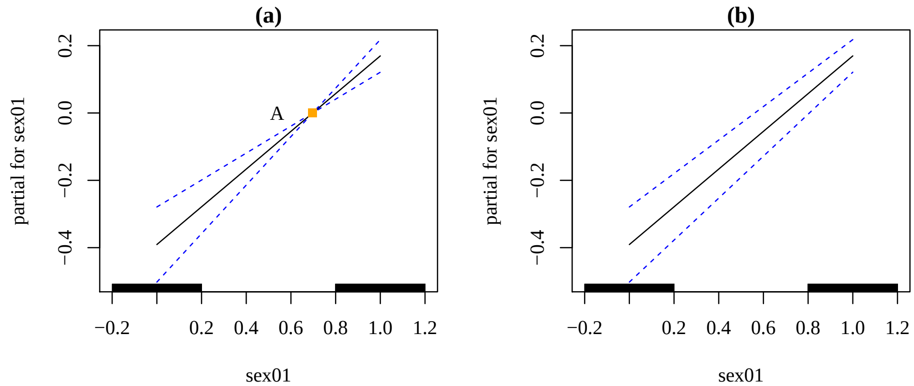


Fig. 8.2 Fitted logistic regression VGAM (the response is `nofriend`) to European-type people in `xs.nz` with (a) `noxmean = FALSE`, (b) `noxmean = TRUE`. In (a), the point A, which has 0 as its SE, has been explictly added to the plot, whereas it has been omitted in (b). The x-coordinate of A is at the mean of the variable `sex01`. The y-coordinate is 0 because component functions are centred.

Chapter 9
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

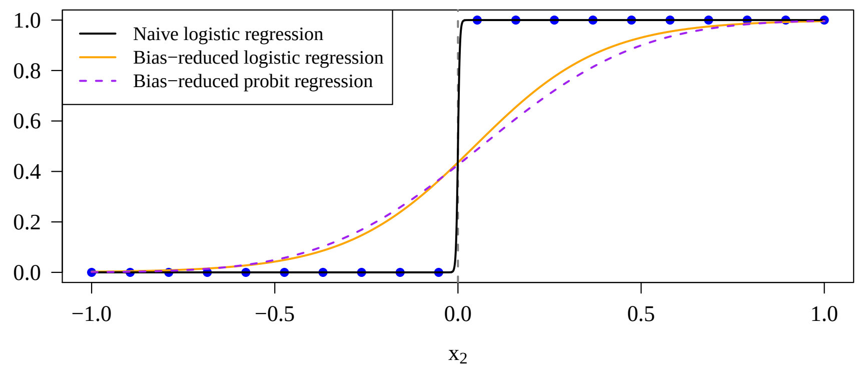


Fig. 9.1 Fitted curves for three binary regression models fitted to completely-separable data, with $n = 20$. A grey vertical line at $x_2 = 0$ is plotted. If there were an additional two points at $(0, 0)$ and $(0, 1)$ then the data would have quasi-complete separation.

Chapter 10
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

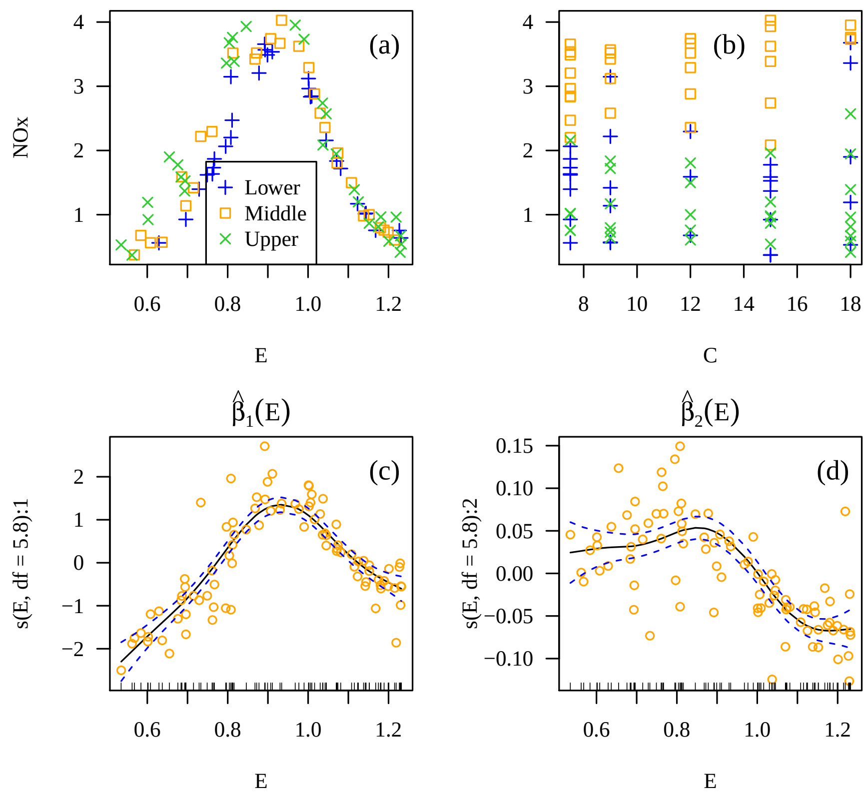


Fig. 10.1 (a)–(b) Scatter plots of the `ethanol` data stratified by `C` and `E` respectively. In (a) `C` has been split into low, medium and high subgroups. In (b) `E` has been split into three similar subgroups—see Fig. 10.2 for an expansion. Plots (c)–(d) are the estimated (centred) component functions of the VCM (10.5) where both the intercept and slope are a smoothing spline of the variable `E`. The dashed lines are pointwise ±2 SEs about the $\widehat{\beta}_j(E)$. The points about the curves are the working residuals.

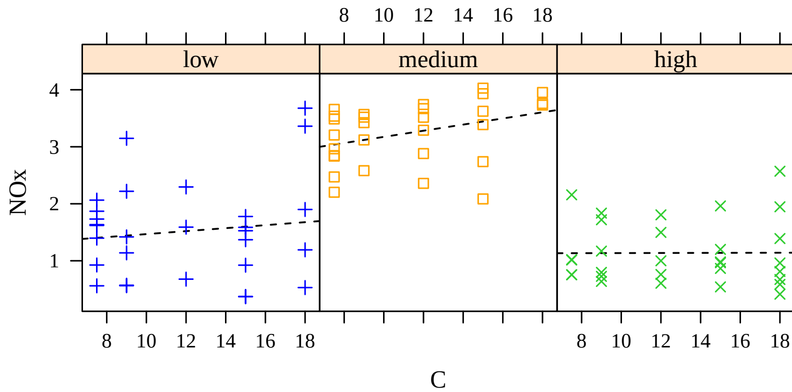


Fig. 10.2 Expansion of Fig. 10.1(b) with a least squares line added to each subset.

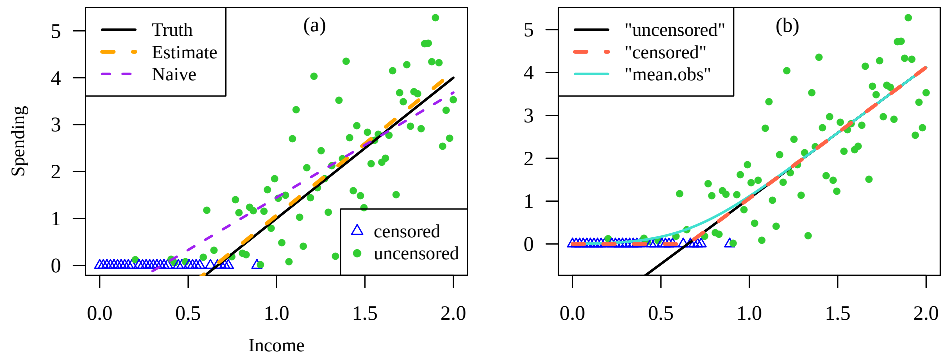


Fig. 10.3 Tobit model fitted to some simulated data. This mimics the type of problem motivating Tobin (1958), viz. spending is non-negative, and is linear beyond a certain income. Values of zero are plotted with a different colour and symbol for clarity. (a) The purple dashed line is a naïve fit that treats all values as if they were 'real'. The estimate and the truth are similar. (b) The 3 types of fitted values currently distinguished by the argument `type.fitted`.

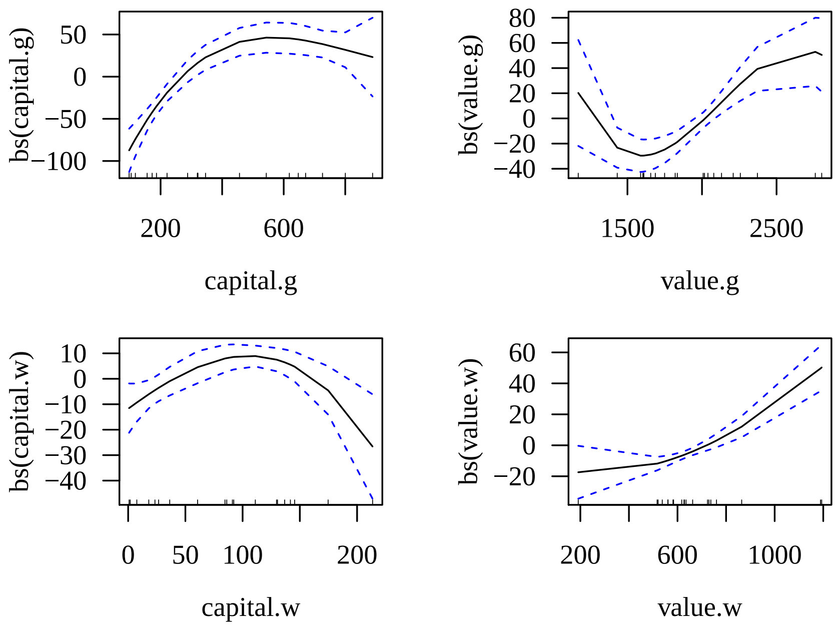


Fig. 10.4 Smooths of a SUR model applied to the `gew` data. The fitted model is called `fit.rs`.

Chapter 11
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

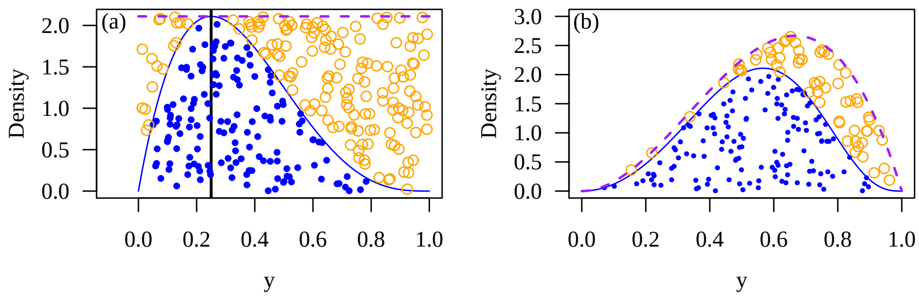


Fig. 11.1 Acceptance-rejection method for generating random variates. The solid blue points are accepted; the hollow orange points are rejected. (a) The density g = `dunif()` is used to generate Beta(2, 4) random variates. The vertical line at $y = 0.25$ denotes the position of the mode, thus defining $C = f(0.25)$. (b) A Kumaraswamy(3, 4) density (f; blue) is overshadowed by a scaled Beta(3, 2) density ($C \cdot g(y)$; purple dashed); the scaling constant is $C = 1.5$.

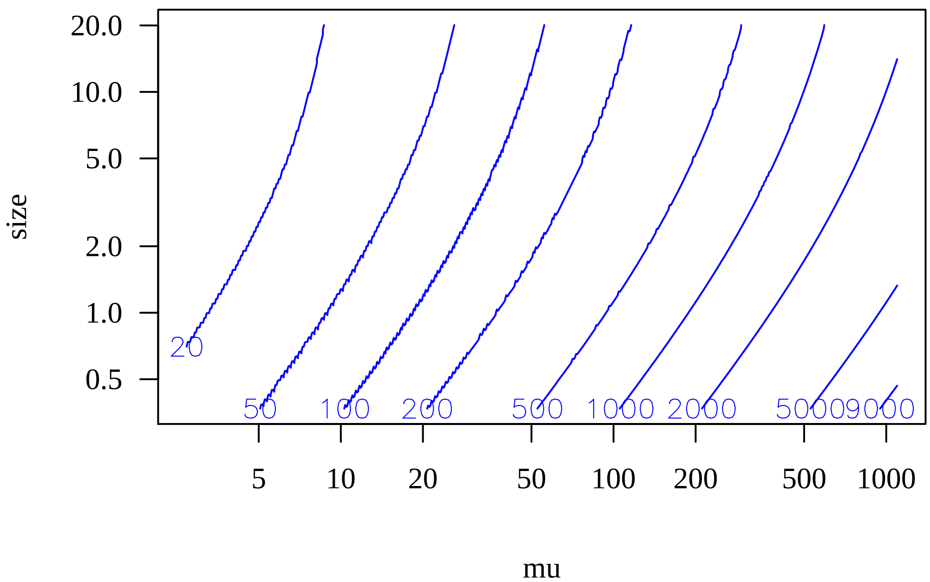


Fig. 11.2 Contour plot of `qnbinom(0.995, size = size, mu = mu)`, where μ and k are on a log-scale. Regions of the (μ, k)-space with a value less than 1000, say, might have the EIM computed by the exact method (11.5). Some of the contour levels appear jagged due to the discrete nature of `qnbinom()`.

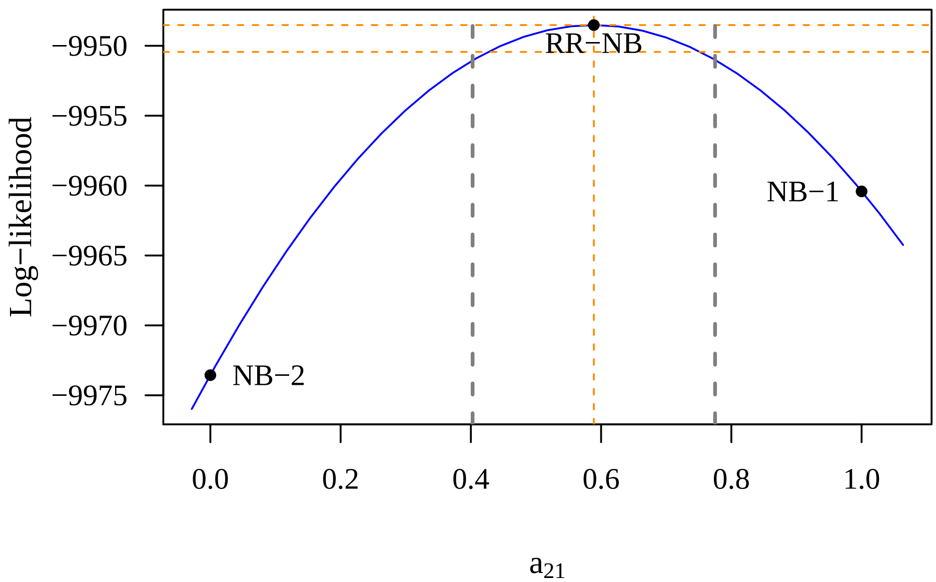


Fig. 11.3 Profile log-likelihood $\ell(a_{21})$ for **rrnb.azpro**; $\ell(\widehat{a}_{21})$ is the highest point. The MLE and likelihood-ratio confidence limits are the orange horizontal lines. The Wald confidence limits are the grey vertical lines. The point at $a_{21} = 0$ is ℓ for NB-2 (intercept-only for k). The point at $a_{21} = 1$ is ℓ for NB-1.

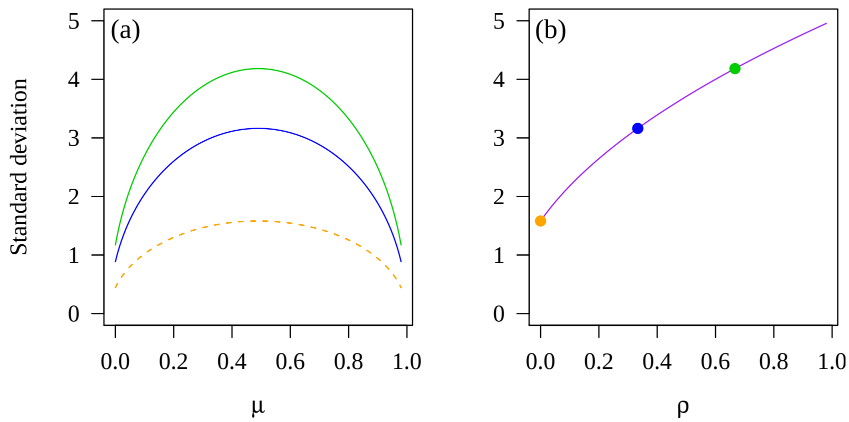


Fig. 11.4 Standard deviation, $\mathrm{SD}(Y_i^*)$, in the beta-binomial distribution, from (11.13) with $N_i = 10$. (a) As a function of μ, for $\rho = 0.33$ (blue), $\rho = 0.67$ (green) and $\rho = 0$ (orange dashed). (b) As a function of ρ, where $\mu_i = 0.5$.

Chapter 12

Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

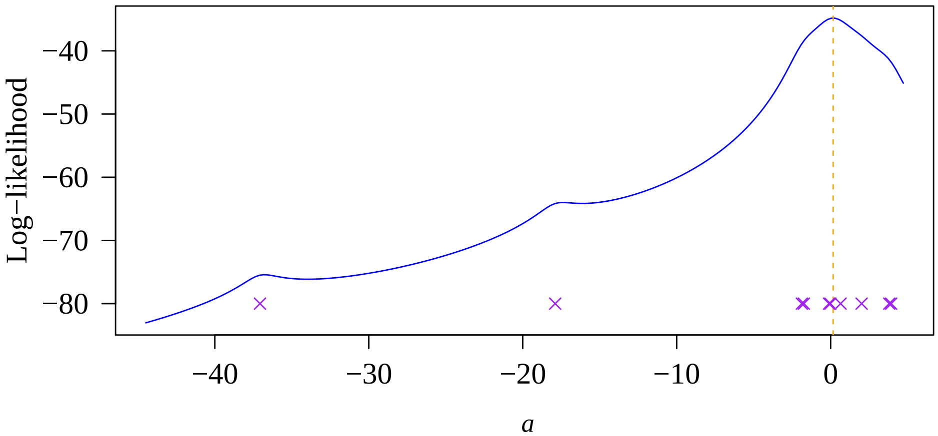


Fig. 12.1 Log-likelihood $\ell(a)$ as a function of the location parameter, for a random sample of size 10 from a standard Cauchy distribution. The dashed vertical line denotes $\widehat{a}$, and the purple $\times$ denotes the data.

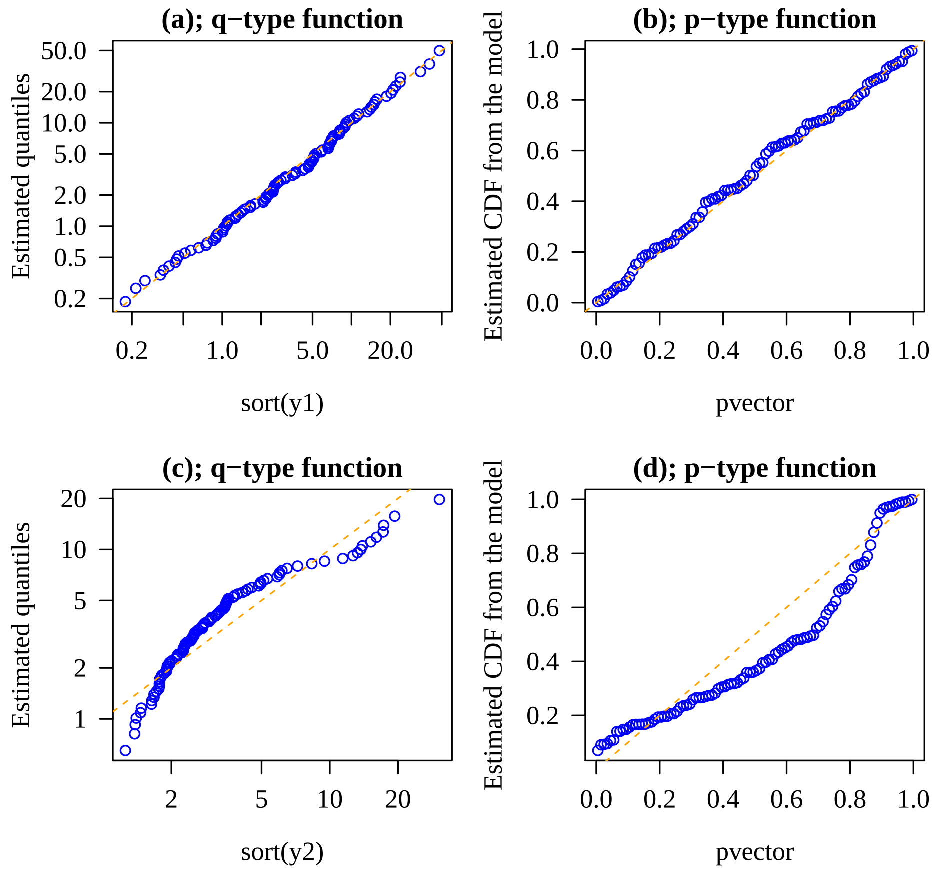


Fig. 12.2 QQ-plots of 2 simulated data sets. A Birnbaum-Saunders distribution is fitted to both. (a)–(b) Data from a Birnbaum-Saunders distribution; (c)–(d) data from a Fréchet distribution. Plots (a) and (c) are based on the `q`-type function, and both axes are on a log-scale. Plots (b) and (d) use the `p`-type function. A $x = y$ line appears in all plots.

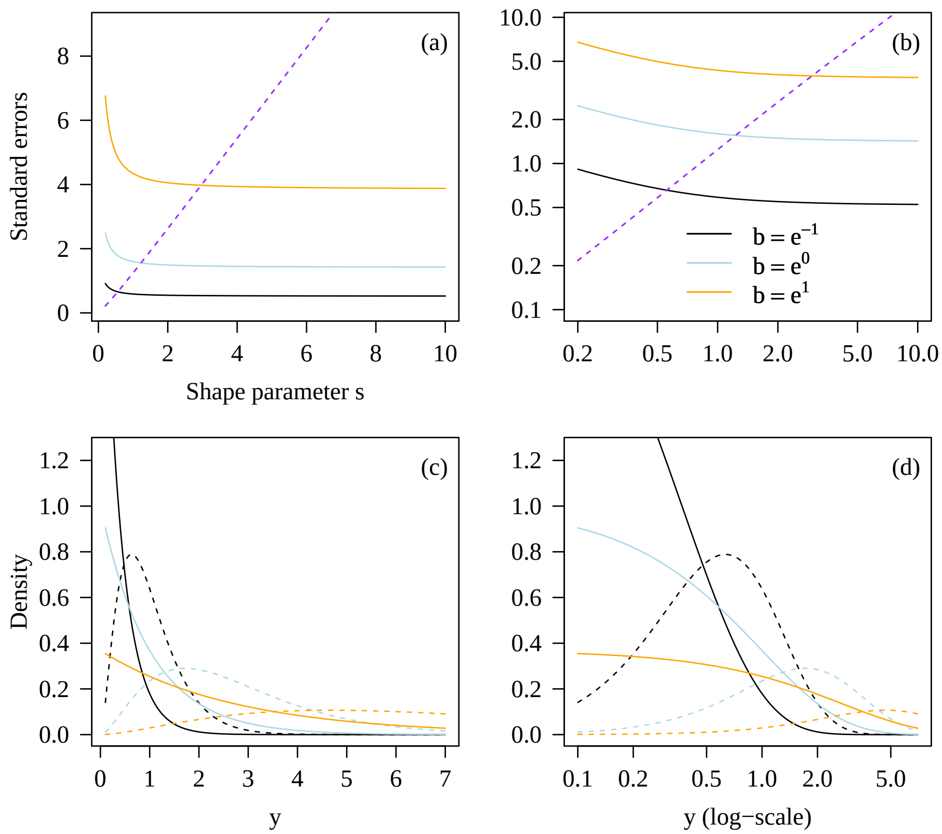


Fig. 12.3 The 2-parameter gamma distribution (12.10). (a) SE($\widehat{b}$) are the solid curves, for various values of the scale parameter b given in the legend of (b). The purple dashed line is SE($\widehat{s}$). (b) The same with both axes on a log-scale. (c) The densities with shape parameters $s = e^0$ (solid lines) and e^1 (dashed). (d) The same as (c) with one axis on a log-scale.

Chapter 13

Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

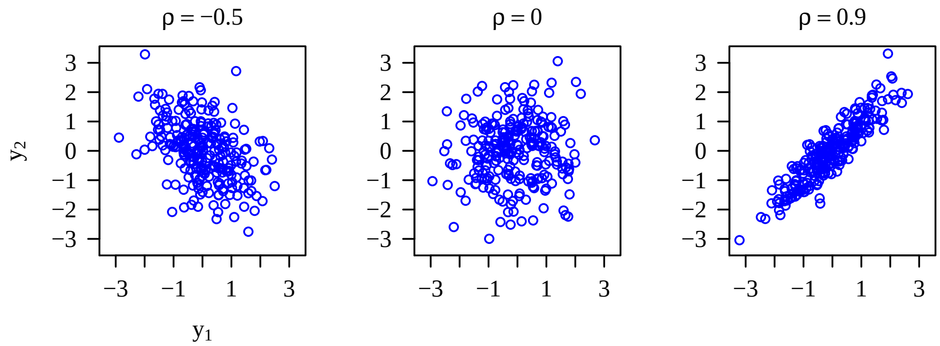


Fig. 13.1 250 random vectors generated from a standard bivariate normal distribution, with various values of ρ.

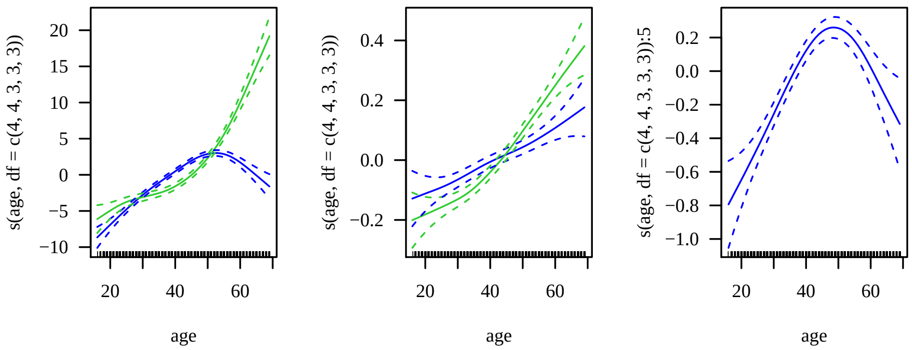


Fig. 13.2 Fitted (centred) component functions of a VGAM N_2 fitted to diastolic and systolic blood pressures data, versus age. The data set are 5649 male Europeans from `xs.nz`. From left to right, those for $\widehat{\mu}_1$ and $\widehat{\mu}_2$, those for $\log \widehat{\sigma}_1$ and $\log \widehat{\sigma}_2$, those for $\log((1+\widehat{\rho})/(1-\widehat{\rho}))$.

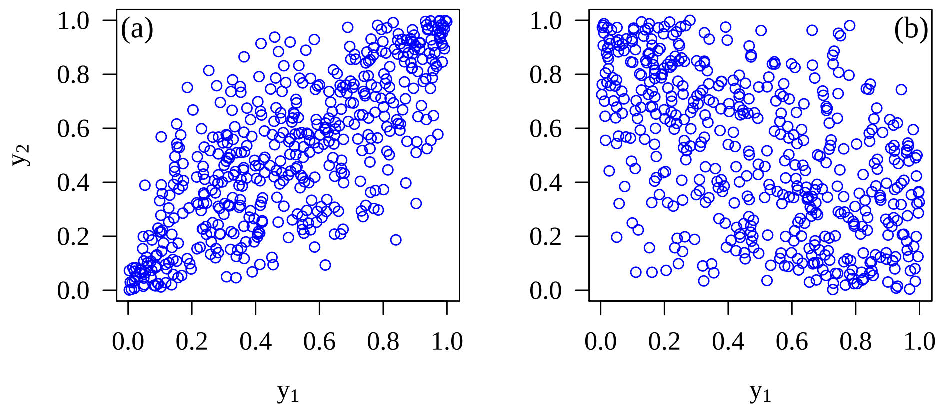


Fig. 13.3 500 random vectors generated from two copulas. (a) Bivariate Gaussian copula with $\alpha = \rho = 0.8$. (b) Bivariate Frank copula with $\alpha = 50$.

Chapter 14

Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

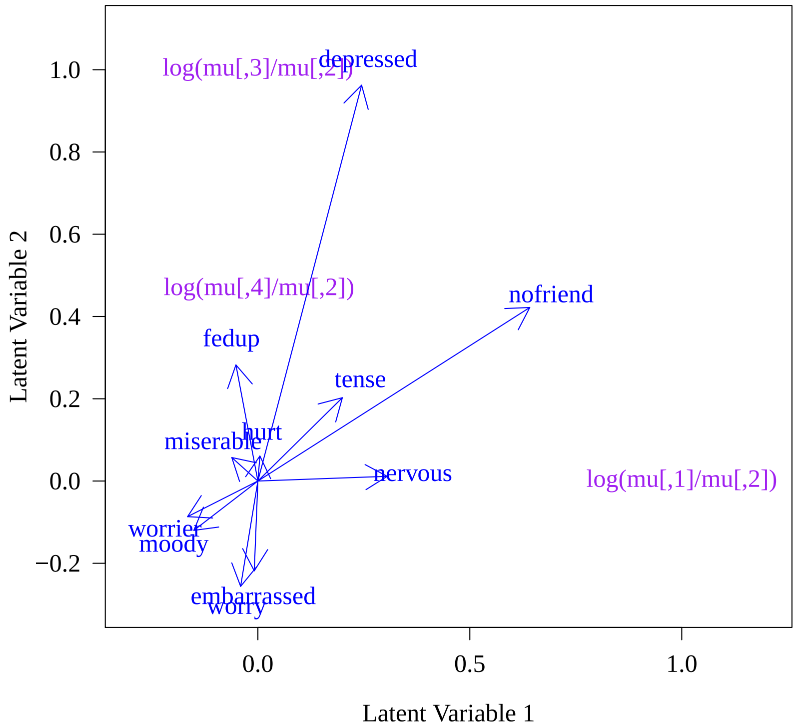


Fig. 14.1 Biplot of a RR-MLM fitted to `marital` status in `xs.nz`, with the latent variables being linear combinations of 11 binary psychological variables. The x- and y-axes are $\widehat{\nu}_1$ and $\widehat{\nu}_2$, respectively. The groups are: 1 single, 2 married/partnered (baseline), 3 divorced/separated, 4 widowed.

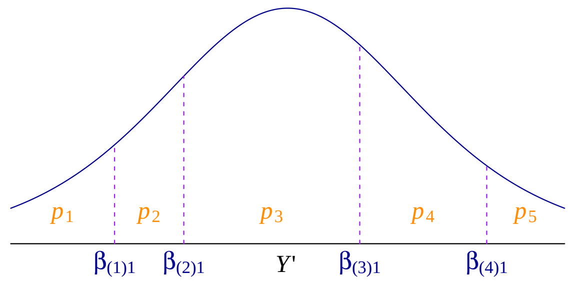


Fig. 14.2 Latent variable Y' interpretation for the cumulative logit model. Here, $\beta_{(j)1}$ are *thresholds* or *cutoff points* so that $\Pr(Y \leq j|\boldsymbol{x}) = \Pr(Y' \leq \beta_{(j)1})$. Then $p_j = \Pr(Y = j|\boldsymbol{x})$. The distribution is a logistic (see (14.18)).

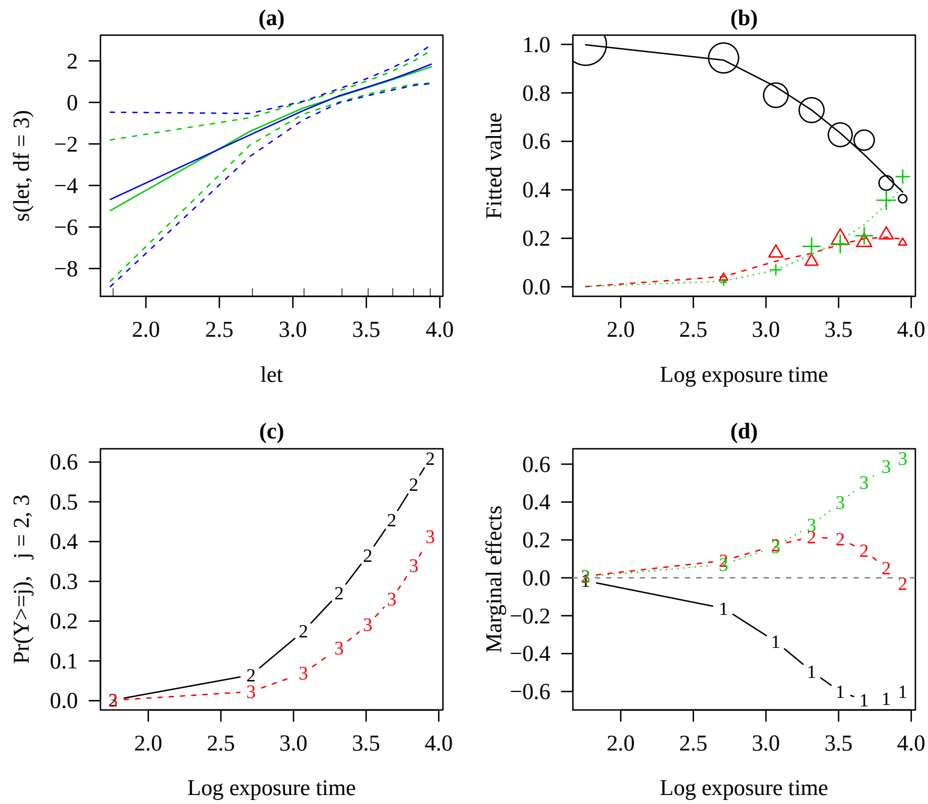


Fig. 14.3 Nonparametric nonparallel cumulative logit model fitted to the `pneumo` data: (a) the two centred functions are overlaid on to one plot; (b) the fitted values as a function of `let` with sample proportions plotted proportional to their counts; (c) reversed cumulative probabilities $\widehat{P}(Y \geq j)$ for $j = 2, 3$. (d) Proportional odds model: the marginal effects for variable `let` for each of the 3 levels.

Chapter 15
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

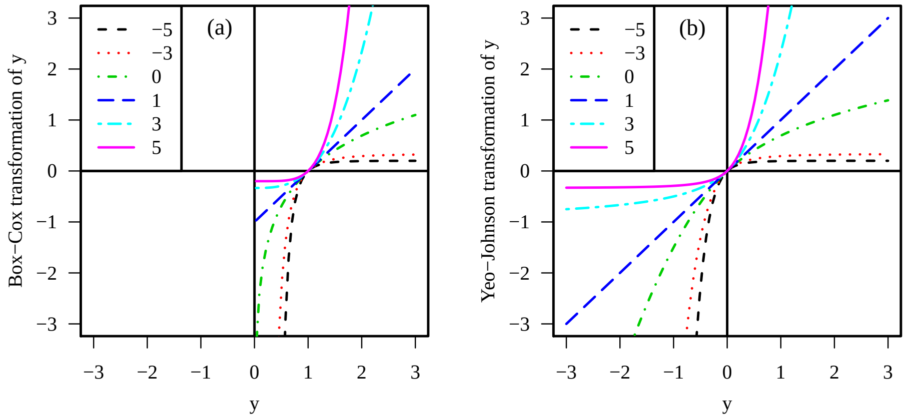


Fig. 15.1 (a) The Box-Cox transformation $(y^{\lambda} - 1)/\lambda$ for various values of λ. (b) The Yeo-Johnson transformation $\psi(\lambda, y)$.

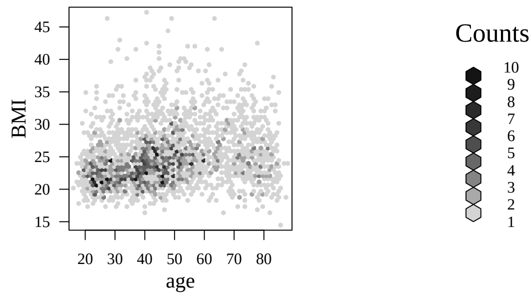


Fig. 15.2 Hexagonal binning plot of `BMI` versus `age` for European-type women in `xs.nz`.

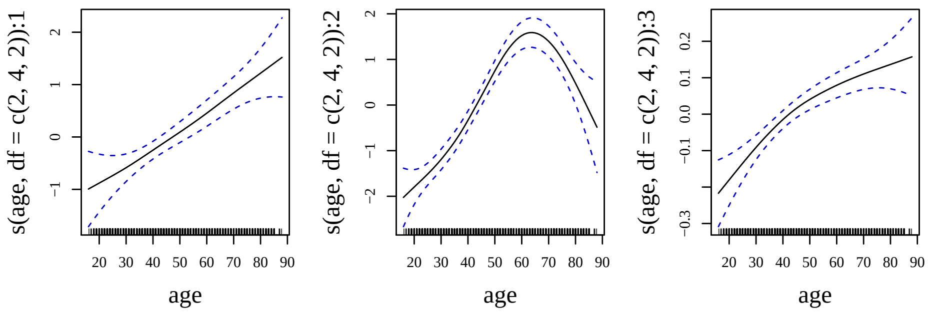


Fig. 15.3 Plot of the VGAM component functions of `w0.LMS0`, a `lms.bcn()` fit of `BMI` versus `age` for European-type women in the `xs.nz` data frame. These are (centred) $\widehat{\lambda}(x)$, $\widehat{\mu}(x)$, $\log\,\widehat{\sigma}(x)$.

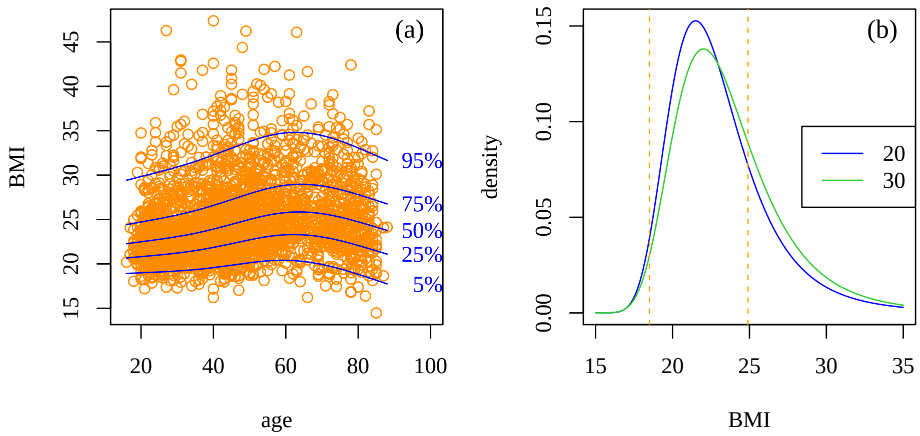


Fig. 15.4 (a) Quantile and (b) density plots of `w0.LMS`. In (b) the estimated densities are for 20 and 30 year olds, and the vertical lines denote an approximate healthy range.

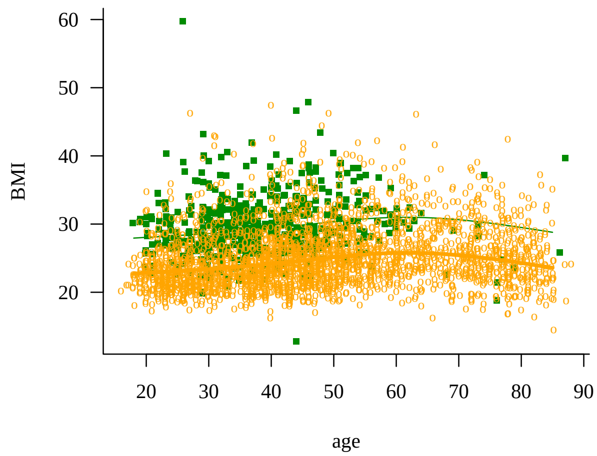


Fig. 15.5 Output of `qtplot()` when there are non-primary variables. The two groups are European and Pacific island women. The fitted 50-percentiles differ by a constant.

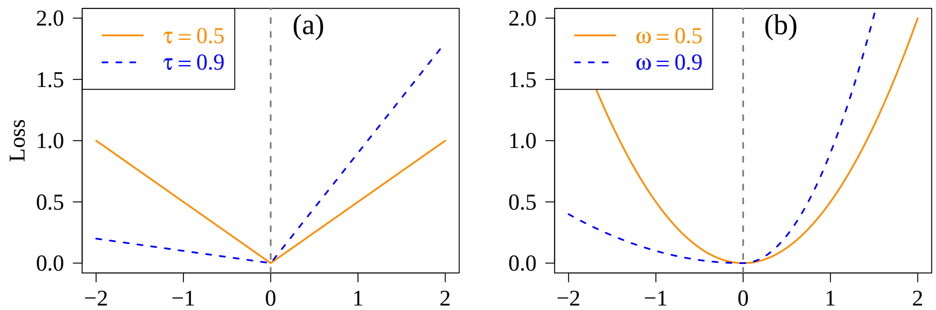


Fig. 15.6 Loss functions for: (a) quantile regression with $\tau = 0.5$ (L_1 regression) and $\tau = 0.9$; (b) expectile regression with $\omega = 0.5$ (least squares) and $\omega = 0.9$. Note: (a) is also known as the *asymmetric absolute loss function* or *pinball loss function*.

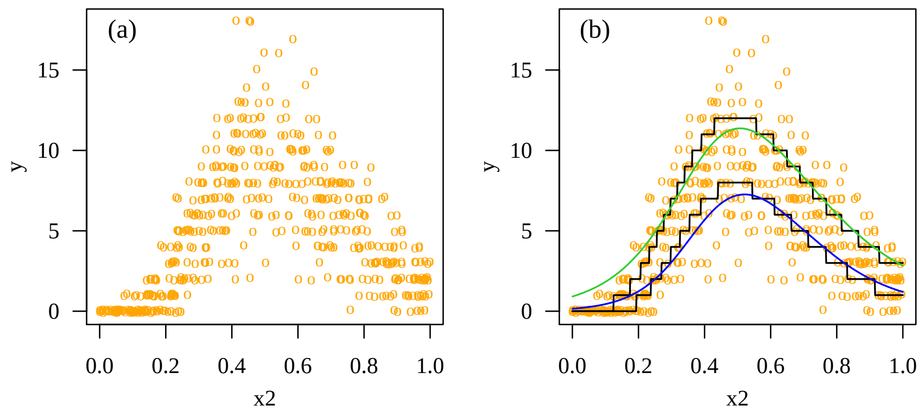


Fig. 15.7 (a) Scatter plot of simulated Poisson counts (15.19). The points have been jittered slightly. (b) Fitted $\mathcal{L}^*(\xi = \exp\{\beta^*_{(s)1} + f^*_{(s)2}(x_{i2})\}$, $\sigma = 1$, $\boldsymbol{\tau} = (\frac{1}{4}, \frac{3}{4})^T)$ model to the data. The smooth curves are the fitted $\tau = 0.25$ and 0.75 quantiles from a vector smoothing spline fit. The step functions are the output from `qpois()` at the corresponding τ values. The actual sample proportions lying below the curves are 35.2 and 79.6 percent. See also Fig. 15.8(b).

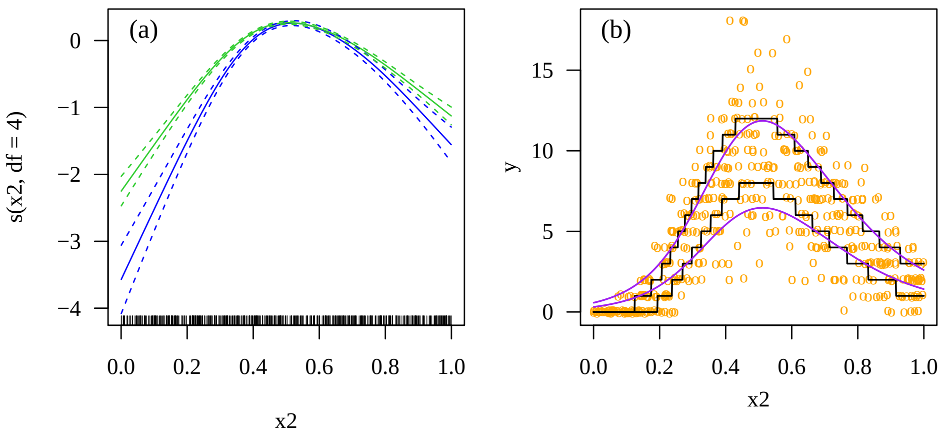


Fig. 15.8 (a) Fitted functions $\widehat{f}_{(s)2}(x_{i2})$ in (15.20) overlaid, with pointwise ± 2 SE bands. (b) Same as Fig. 15.7(b) except the quantiles (purple curves) are constrained to be parallel on the η-scale (log scale here). The actual sample proportion lying below the curves are 36.4 and 78.4 percent.

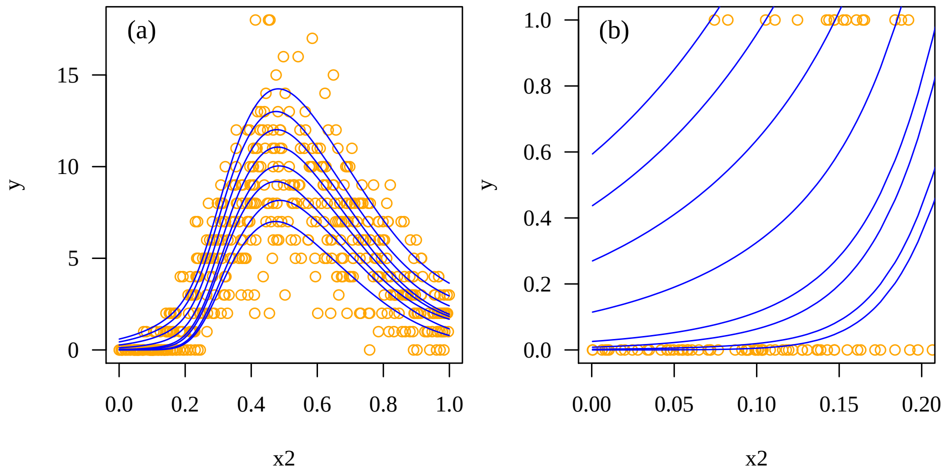


Fig. 15.9 (a) Quantile plot from the onion method applied to some simulated Poisson data. The quantiles are positive, nonparallel and noncrossing. Here, $\boldsymbol{\tau}$ has values 0.2(0.1)0.9. (b) A zoom-in view of the LHS of (a).

Then Fig. 15.9(a) was produced by

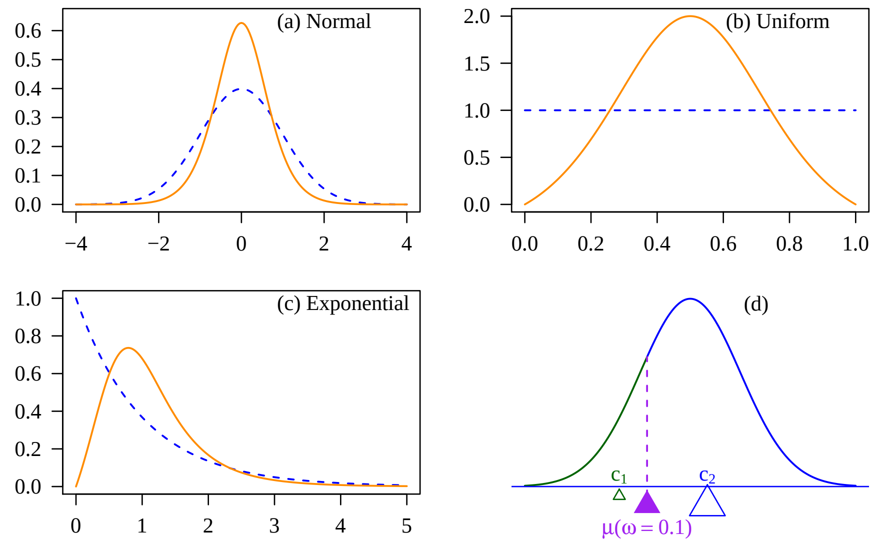


Fig. 15.10 (a)–(c) Density plots of expectile derived g (orange solid lines; (15.30)) for the original f of standard normal, standard uniform and standard exponential distributions (blue dashed lines). (d) Illustration of the interpretation of expectiles in terms of centres of balance, the hollow triangles at positions c_1 and c_2. Here, the vertical dashed line is at the 0.1-expectile, the solid triangle at $\mu(\omega = 0.1)$, which means that (15.33) is satisfied with $\omega = 0.1$.

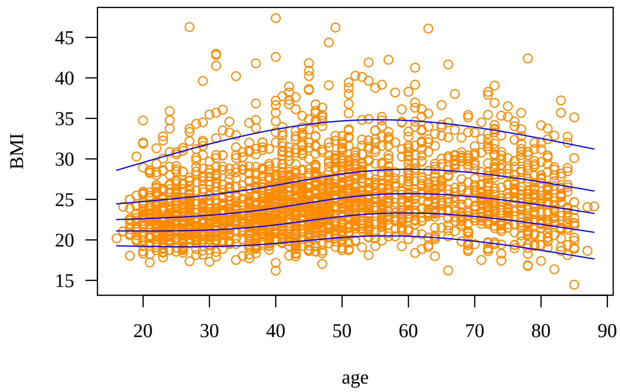


Fig. 15.11 'Quantile' plot from `amlnormal()` applied to `women.eth0`: 5, 25, 50, 75, 95 'percentile' curves. Each regression curve is a regression spline with 3 degrees of freedom (1 = linear fit).

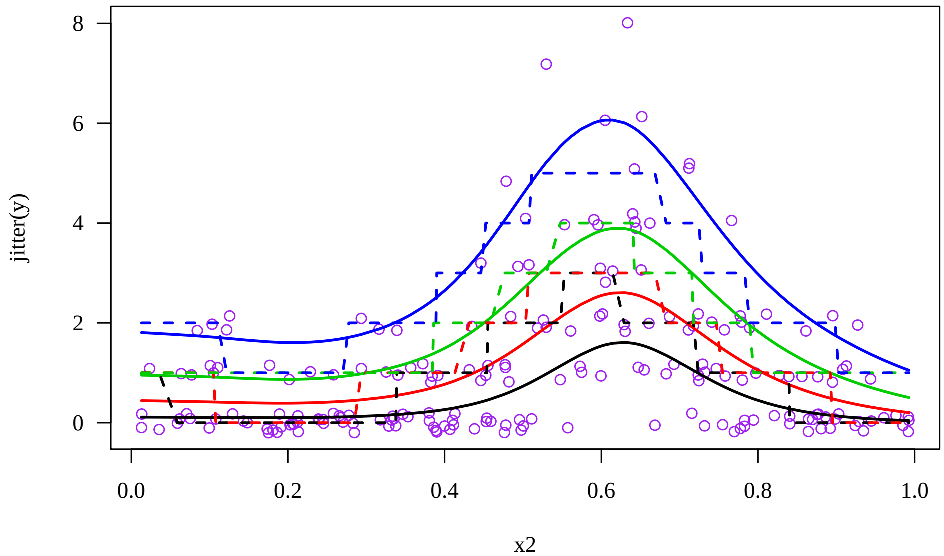


Fig. 15.12 Quantile plot from an `amlpoisson()` fit using simulated data. The dashed step functions are the `qpois()` quantiles calibrated with the true mean function and the proportion under the expectile curves. The response has been jitted to aid clarity.

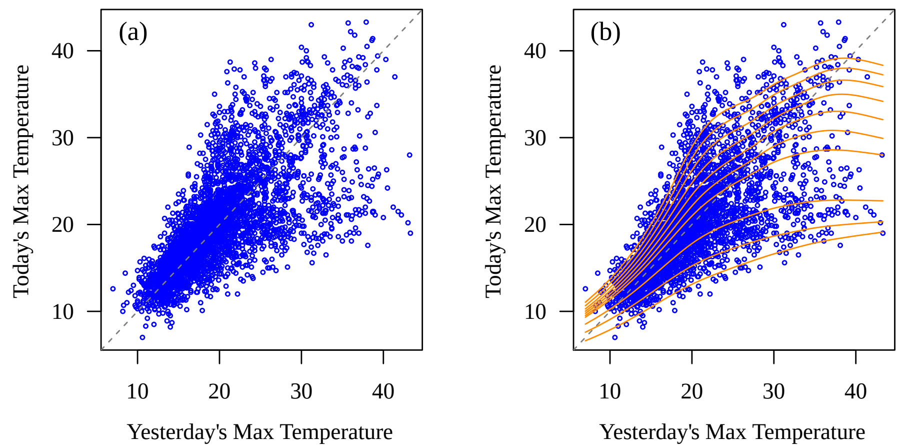


Fig. 15.13 (a) Melbourne maximum temperature data, called `melbmaxtemp`, in °C. (b) Onion method using `amlnormal()` expectiles applied to the data with an assortment of w_j values. The sample proportions below the curves are 1, 4.8, 20.2, 54.9, 65.2, 74.8, 82.7, 88.2, 91.7, 94.8 percent.

Chapter 16
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R

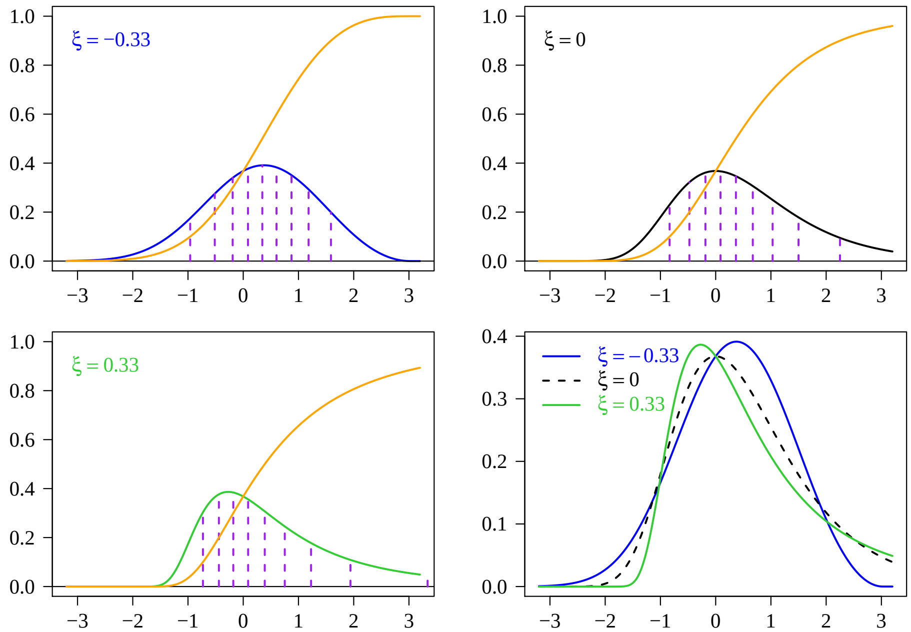


Fig. 16.1 GEV densities for values $\mu = 0$, $\sigma = 1$, and $\xi = -\frac{1}{3}$, 0, $\frac{1}{3}$ (Weibull-, Gumbel- and Fréchet-types respectively). The orange curve is the CDF, the dashed purple segments divide the density into areas of $\frac{1}{10}$. The bottom RHS plot has the densities overlaid.

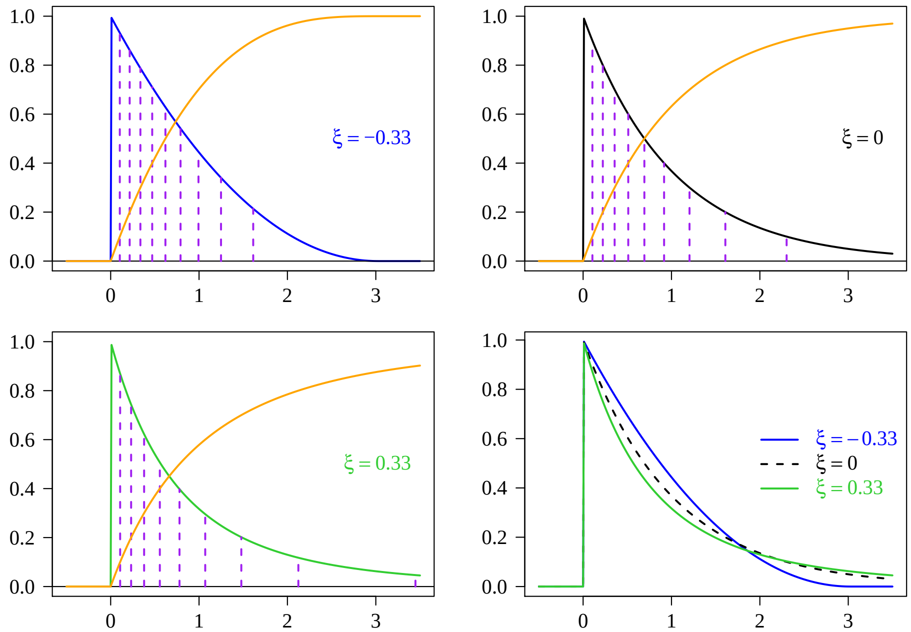


Fig. 16.2 GPD densities for values $\mu = 0$, $\sigma = 1$, and $\xi = -\frac{1}{3}$, 0, $\frac{1}{3}$ (beta-, exponential- and Pareto-types, respectively). The orange curve is the CDF, the dashed purple segments divide the density into areas of $\frac{1}{10}$. The bottom RHS plot has the densities overlaid.

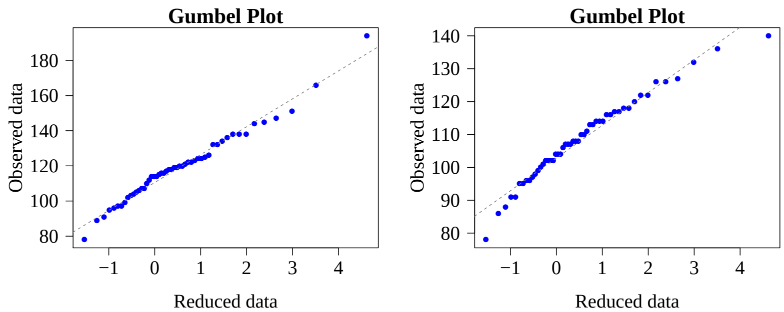


Fig. 16.3 Gumbel plot of the two highest annual sea levels of the Venice data (`venice`).

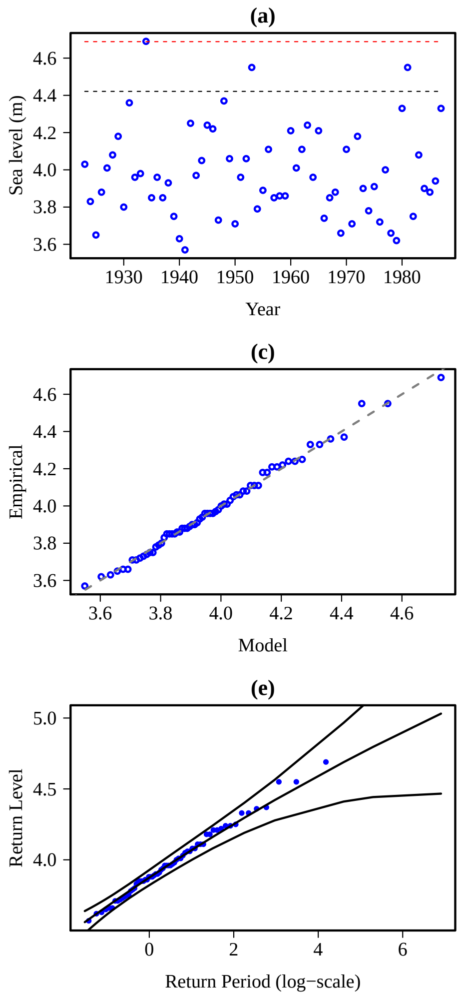


Fig. 16.4 Intercept-only GEV model fitted to the `portpirie` annual maximum sea levels data. (a) Scatter plot, and the dashed horizontal lines are the resulting 95% and 99% quantiles. (b) Probability plot. (c) Quantile plot. (d) Density plot. (e)–(f) Return level plots, with slight changes in the x-axis labelling.

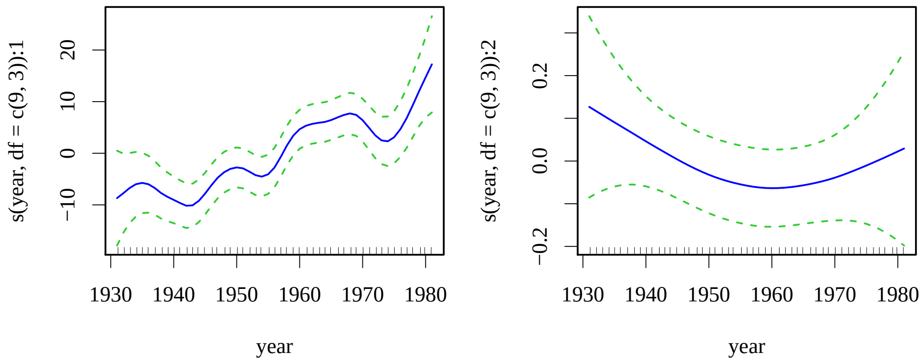


Fig. 16.5 VGAM fitted to the Venice sea level data (`fit1`).

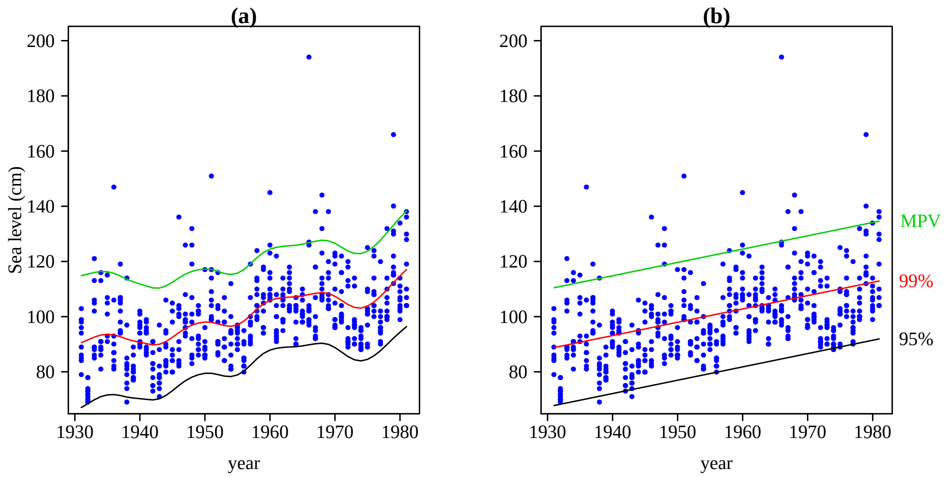


Fig. 16.6 (a) Quantile plot of the Venice sea level data (`fit2`). (b) Venice data overlaid with the fitted values of `fit3`. This model underfits.

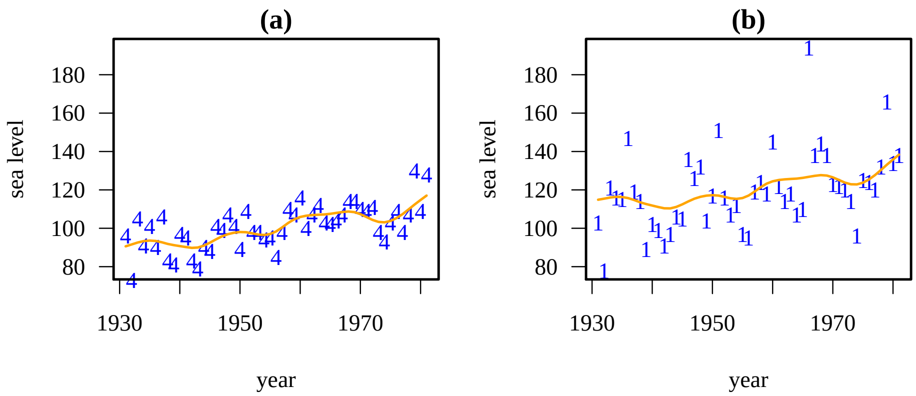


Fig. 16.7 (a) The fitted 99 percentile of `fit2` and the 99 percentile data values (4th highest sea level each year). (b) Fitted median predicted value (MPV) of the Venice sea level data from `fit2`. The points are the highest sea levels for each year.

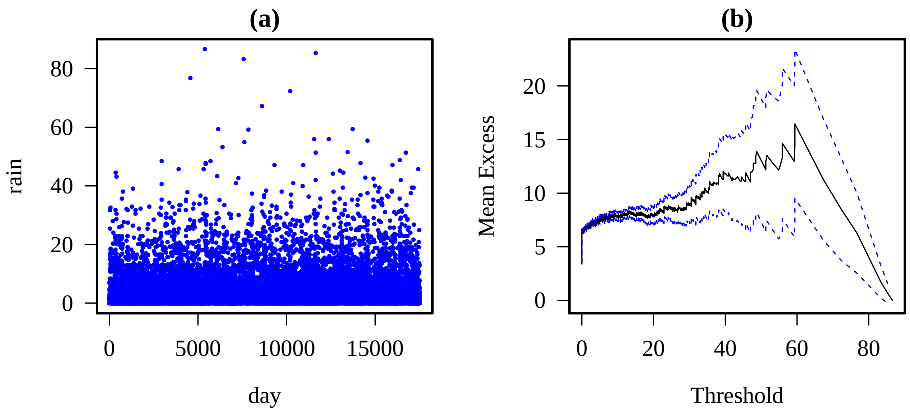


Fig. 16.8 The `rain` daily rainfall data. (a) Scatter plot. (b) Mean excess plot.

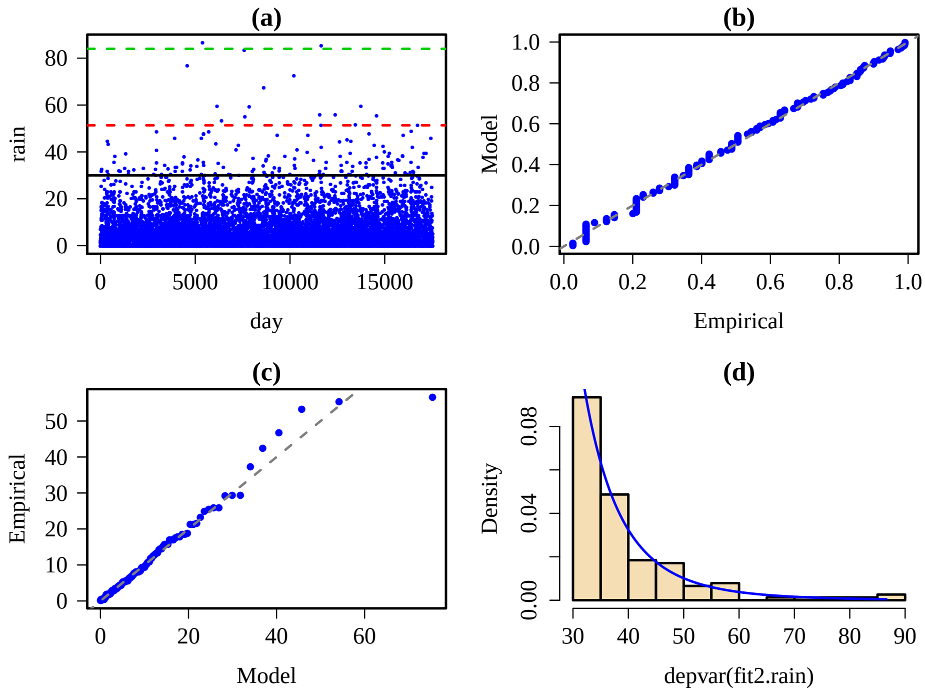


Fig. 16.9 Intercept-only GPD model fitted to the `rain` daily rainfall data. (a) Scatter plot, with the solid black horizontal line denoting the threshold at 30. The dashed horizontal lines are the resulting 90% and 99% quantiles. (b) Probability plot. (c) Quantile plot. (d) Density plot.

Chapter 17
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R
© T. W. Yee, 2015

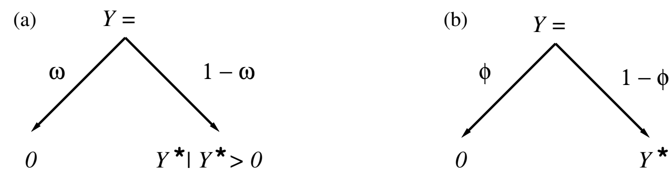


Fig. 17.1 Decision tree diagram for (a) zero-alteration versus (b) zero-inflation. They depict (17.4) and (17.7) respectively. Here, Y^* corresponds to a parent distribution such as the binomial or Poisson, and Y is the response of interest. The probabilities ω and ϕ dictate the decisions.

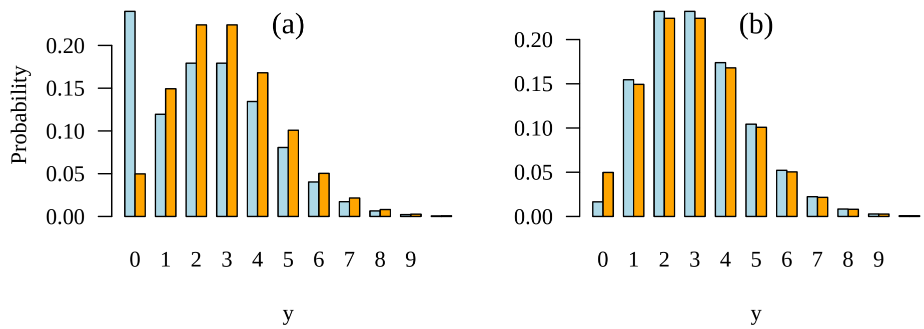


Fig. 17.2 Probability functions of a (a) zero-inflated Poisson with $\phi = 0.2$, (b) zero-deflated Poisson with $\omega = -0.035$. Both are compared to their parent distribution which is a Poisson($\mu = 3$) in orange.

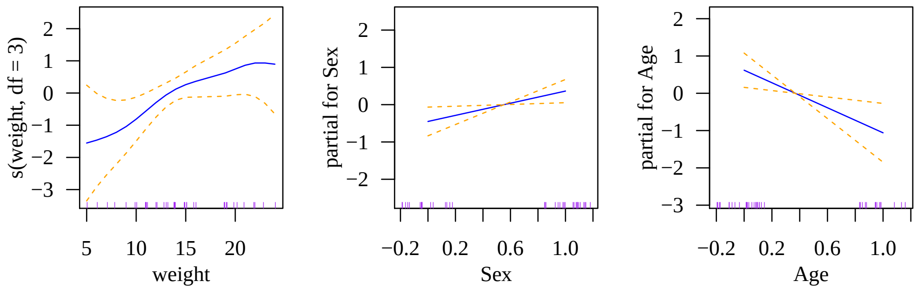


Fig. 17.3 Estimated component functions for the `deermice` VGAM.

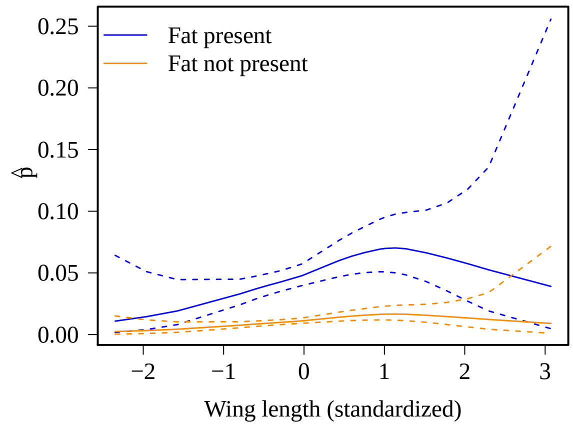


Fig. 17.4 Capture probability estimates with approximate ±2 pointwise SEs, versus wing length with (blue) and without (orange) fat content present fitting a $\mathcal{M}_h$-VGAM, using the `prinia` data.

Appendix A
Figures from Vector Generalized Linear and Additive Models: With an Implementation in R
© T. W. Yee, 2015

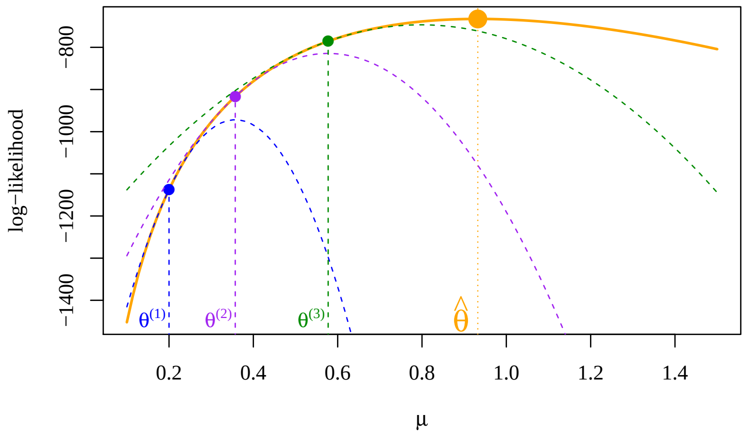


Fig. A.1 The first few Newton-like iterations for a Poisson regression fitted to the `V1` data set. The solid orange curve is $\ell(\theta)$ with $\theta = \mu$. The initial value is $\theta^{(1)} = 0.2$. Each iteration $\theta^{(a)}$ corresponds to the maximum of the quadratic (dashed curves) from the previous iteration.

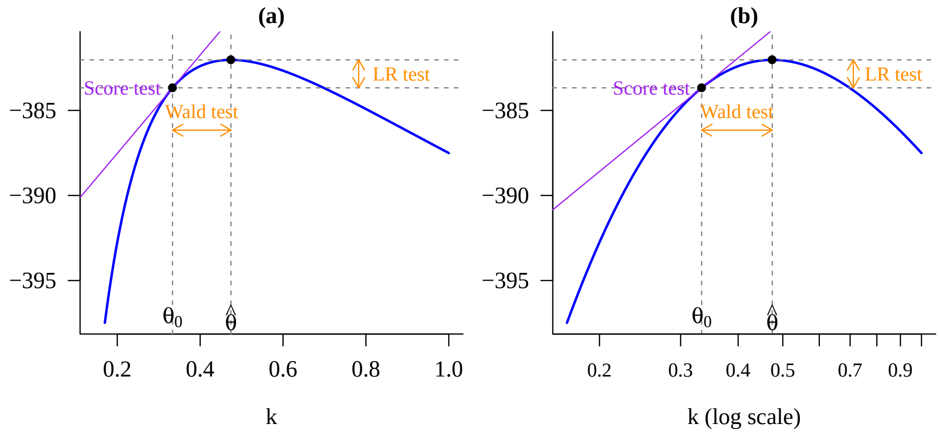


Fig. A.2 Negative binomial NB(μ, k) distribution fitted to the `machinists` data set. The y-axis is ℓ. Let $\theta = k$ and $\theta^* = \log\, k$. (a) $\ell(\theta)$ is the solid blue curve. (b) $\ell(\theta^*)$ is the solid blue curve. Note: for $H_0 : \theta = \theta_0$ (where $\theta_0 = \frac{1}{3}$), the likelihood-ratio test, score test and Wald test statistics are based on quantities highlighted with respect to ℓ. In particular, the score statistic is based on the tangent $\ell'(\theta_0)$.

References

Yee, T. W. (2015). *Vector Generalized Linear and Additive Models: With an Implementation in R.* New York, USA: Springer.