University of Zurich[UZH]

EBPI Epidemiology, Biostatistics and Prevention Institute

stats::free1way(): Semiparametrically Efficient Population and Permutation Inference in Distribution-free Stratified K-sample Oneway Layouts

Torsten Hothorn & Kurt Hornik

useR! 2026, Warsaw

Neck Pain Reduction by Laser Therapy

Pain 124 (2006) 201–210

PAIN

www.elsevier.com/locate/pain

The effect of 300 mW, 830 nm laser on chronic neck pain: A double-blind, randomized, placebo-controlled study

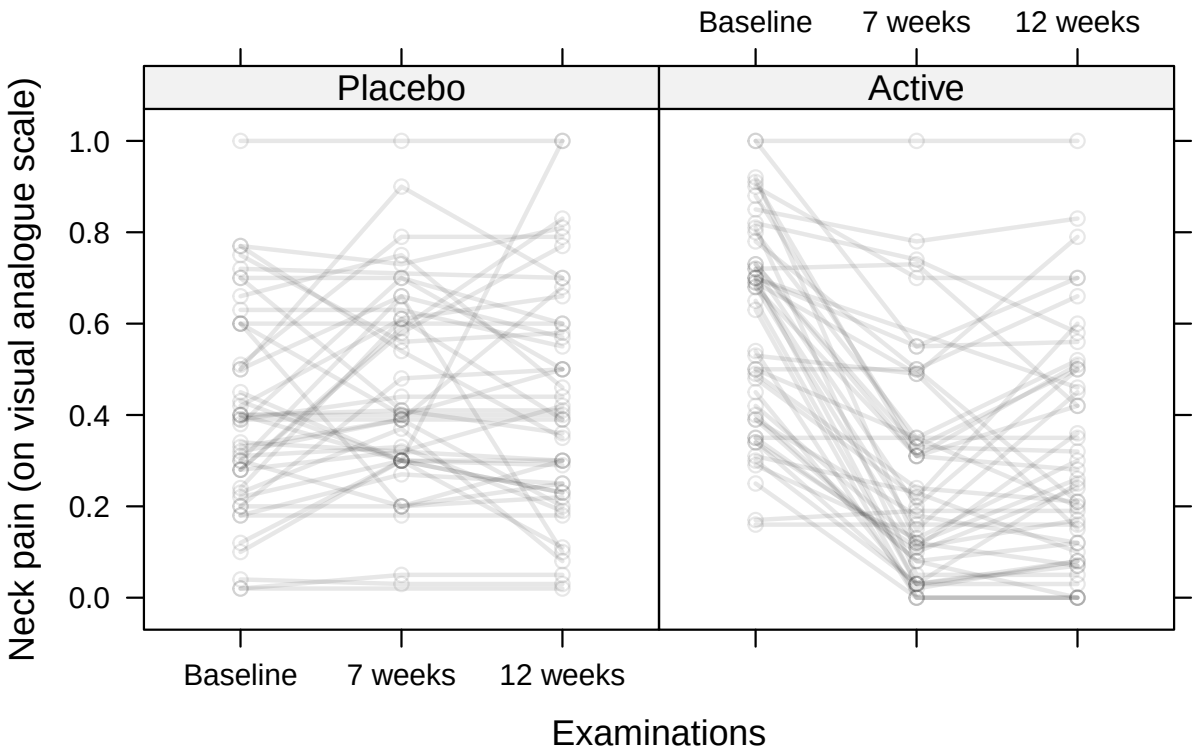
Roberta T. Chow [a,*], Gillian Z. Heller [c], Les Barnsley [b]

[a] *Castle Hill Medical Centre, 269-271 Old Northern Road, Castle Hill, NSW 2154, Australia*
[b] *Department of Statistics, Macquarie University, NSW 2109, Australia*
[c] *Faculty of Medicine, University of Sydney, Department of Rheumatology, Repatriation Hospital, Concord, NSW 2139, Australia*

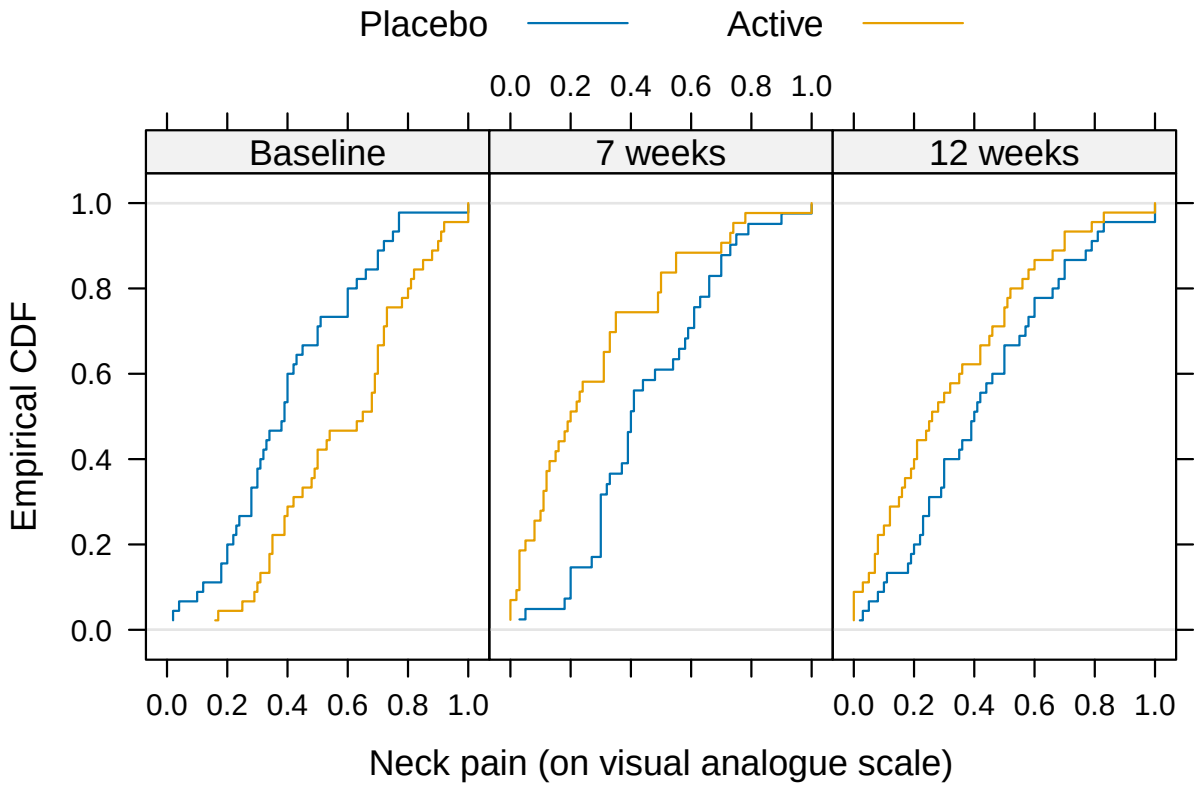
Received 16 June 2005; received in revised form 23 March 2006; accepted 17 May 2006

Contains a discussion about how to analyse VAS and motivated research accumulating in **ordinalCont**

Neck Pain Reduction by Laser Therapy: Joint



Neck Pain Reduction by Laser Therapy: Marginal



stats::wilcox.test()

```
> wilcox.test(vas ~ laser, data = pain,
+             subset = time == "12 weeks",
+             exact = TRUE, conf.int = TRUE)

        Wilcoxon rank sum exact test

data:  vas by laser
W = 1260, p-value = 0.046
alternative hypothesis: true location shift is not equal to 0
96 percent confidence interval:
 0.00 0.22
sample estimates:
difference in location
                  0.11
```

Wilcoxon 1945

INDIVIDUAL COMPARISONS BY RANKING METHODS

FRANK WILCOXON
American Cyanamid Co.

For 12-week ranked VAS measurements $R_i = \text{rank}(Y_i)$ with $i = 1, \ldots, n$ for placebo and $i = n + 1, \ldots, N$ for active laser, the rank sum statistic is

$$\sum_{i=1}^{n} R_i \propto \underbrace{\frac{1}{N-n} \sum_{i=n+1}^{N} R_i - \frac{1}{n} \sum_{i=1}^{n} R_i}_{t\text{-test on ranks}}$$

testing the CDF hypothesis

$$H_0 : F(y \mid \text{placebo}) = F(y \mid \text{active}) \quad \forall y \in (0, 1)$$

Bauer 1972

© Journal of the American Statistical Association
September 1972, Volume 67, Number 339
Theory and Methods Section

Constructing Confidence Sets Using Rank Statistics

DAVID F. BAUER*

$$F(y \mid \text{active}) = F(y - \beta \mid \text{placebo})$$

Hodges-Lehmann estimate and confidence interval for location-shift

Problems

- location shift inappropriate for bounded or discrete distributions
- test not necessarily optimal against such alternatives
- model diagnostics?
- `power.wilcox.test()`?
- stratification?

Lehmann 1953

THE POWER OF RANK TESTS[1]

BY E. L. LEHMANN

Stanford University and University of California, Berkeley

$$F(y \mid \text{active}) = m(F(y \mid \text{placebo}), \delta)$$

maybe $m(p, \delta) = p^{\delta}$ ("Lehmann alternative")?

Solutions

STATISTICS IN MEDICINE, VOL. 12, 2257–2271 (1993)

SAMPLE SIZE CALCULATIONS FOR ORDERED CATEGORICAL DATA

JOHN WHITEHEAD
Department of Applied Statistics, University of Reading, Earley Gate, P.O. Box 238, Reading RG6 2AL, U.K.

$$m(p, \delta) = \text{expit}(\text{logit}(p) - \delta)$$

aka “proportional odds (PO) model” with δ a log-odds ratio

NB: `wilcox.test()` is the nonparametric (or empirical) score test in a PO model for binary, ordered, continuous outcomes.

Distribution-free Models in Oneway Layouts

For

- strata (S, for $b = 1, \ldots, B$),
- treatment groups (G, for $k = 2, \ldots, K$), and
- outcome values y based on
- effect parameters $\delta_2, \ldots, \delta_K$:

$$\underbrace{F(y \mid G = k, S = b)}_{\text{Treatment } k \text{ CDF in } b} = m(\underbrace{F(y \mid G = 1, S = b)}_{\text{Placebo CDF in } b, \text{unspecified}}, \underbrace{\delta_k}_{k \text{ vs. Placebo}})$$

`stats::free1way()`, termed after SAS' `NPAR1WAY`, since R 4.6.0 for semiparametrically efficient population and permutation inference in distribution-free stratified K-sample oneway layouts.

Implementation in stats

`Enhances: wilcox.test()`, `kruskal.test()`, `friedman.test()`, `mantelhaen.test()`, `prop.test()`, `mcnemar.test()`, `MASS::polr()`, `tram::Polr`, and (as a role model) `rms::orm()`, ...

Fast joint optimisation of empirical likelihood wrt effect and nuisance parameters by Newton-Raphson with sparse analytic Hessians, courtesy by Frank Harrell.

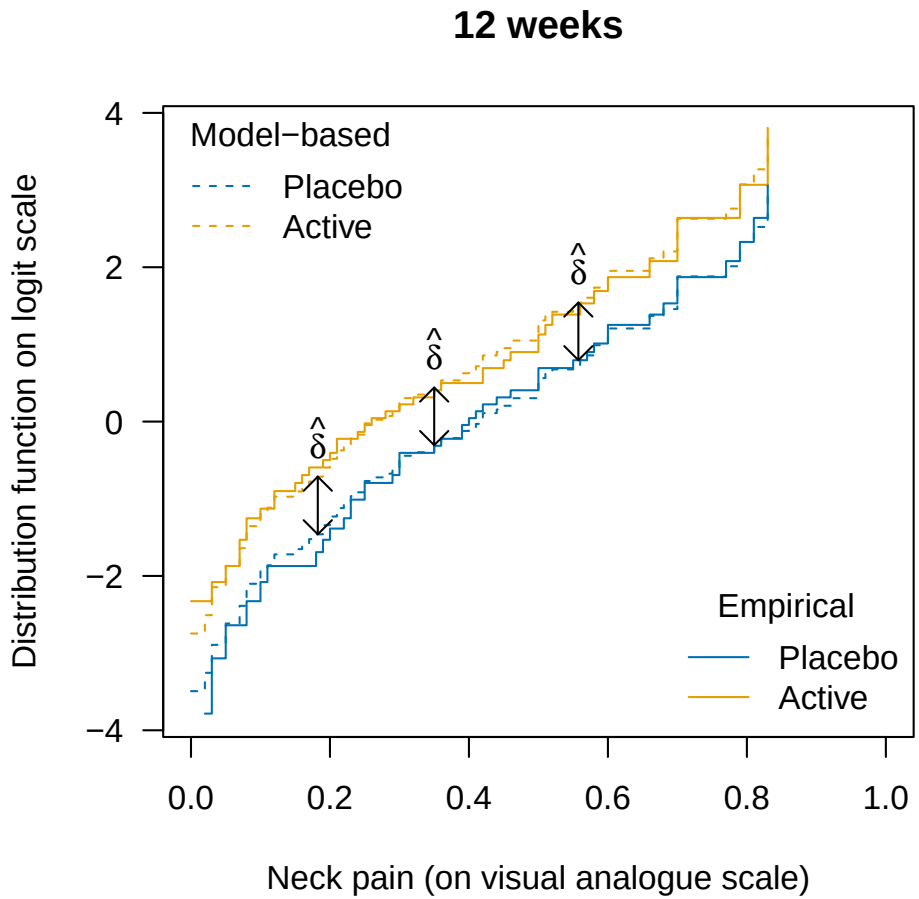
Applicable to binary, ordinal, discrete, continuous, and potentially (independently) right-censored observations.

Special Models

- $m_{\mathsf{L}}(p, \delta) = p^{\exp(-\delta)}$ (Lehmann alternatives, loglog link),
- $m_{\mathsf{PH}}(p, \delta) = 1 - (1 - p)^{\exp(-\delta)}$ (proportional hazards, cloglog link)
- $m_{\mathsf{PO}}(p, \delta) = \text{expit}(\text{logit}(p) - \delta)$ (proportional odds, logit link), and
- $m_{\mathsf{Cd}}(p, \delta) = \Phi(\Phi^{-1}(p) - \delta)$ (generalised Cohen's d, probit link)

Against these alternatives, tests implemented in `free1way()` are locally most powerful.

Pain Reduction: Model Diagnostics



Pain Reduction: Parameter Estimation

```
> f <- free1way(vas ~ laser, data = pain,
+               subset = time == "12 weeks",
+               exact = TRUE)
> ### ### empirical log-likelihood
> logLik(f)
[1] -338.214
> ### estimated log-odds ratio + SE
> c(coef(f), sqrt(vcov(f)))
laserActive
  -0.746923    0.372346
```

Pain Reduction: Permutation Score Inference

```
> summary(f, test = "Permutation")

        2-sample Wilcoxon test against proportional odds
        alternatives

data:  vas by laser (Placebo, Active)
W = -2.015, p-value = 0.046
alternative hypothesis: true log-odds ratio is not equal to 0
> confint(f, test = "Permutation")
                 2.5 %     97.5 %
laserActive -1.48199 -0.0119882
```

Pain Reduction: Empirical Likelihood Inference

```
> summary(f, test = "Rao")

        2-sample Wilcoxon test against proportional odds
        alternatives

data:  vas by laser (Placebo, Active)
Rao chi-squared = 4.062, df = 1, p-value = 0.0439
alternative hypothesis: true log-odds ratio is not equal to 0
> confint(f, test = "Rao")
                 2.5 %     97.5 %
laserActive -1.47361 -0.0203546
```

`test = "LRT"` and `test = "Wald"` also available.

Pain Reduction: Power Assessment

```
> power.free1way.test(delta = coef(f),
+                     power = .8, sig.level = .05)

      2-sample Wilcoxon test against proportional odds
      alternatives

                  n = 88
Total sample size = 88 (Control) + 88 (laserActive) = 176
            power = 0.801078
        sig.level = 0.05
   Standard error = 0.266241
   log-odds ratio = -0.746923

NOTE: 'n' is sample size in control group
```

Pain Reduction: Repeated Measurements

patient = block, aka `friedman.test()` in each group:

```
> pain$tc <- with(pain, interaction(laser, time))
> levels(pain$tc)[1:2] <- "Baseline"
> coef(m <- free1way(vas ~ tc | id, data = pain))
 tcPlacebo.7 weeks   tcActive.7 weeks tcPlacebo.12 weeks
          1.144270          -5.623352           0.407917
 tcActive.12 weeks
         -4.352816
```

Pain Reduction: Repeated Measurements

```
> library("multcomp")
> K <- matrix(c(-1, 1,  0, 0,
+                0, 0, -1, 1), byrow = TRUE, nrow = 2)
> rownames(K) <- c("7 weeks: Active - Placebo",
+                  "12 weeks: Active - Placebo")
> confint(glht(m, linfct = K))$confint
                             Estimate      lwr      upr
7 weeks: Active - Placebo  -6.76762 -8.72808 -4.80716
12 weeks: Active - Placebo -4.76073 -6.57525 -2.94622
attr(,"conf.level")
[1] 0.95
attr(,"calpha")
[1] 2.18344
```

Resources

Visit

- `stats::free1way`
- `stats::rfree1way`
- `stats::power.free1way.test`
- `stats::ppplot`

and add-on package **free1way.docreg** at `CRAN` near you:

- all the models
- likelihood and inference procedures
- optimisation algorithm
- literate pure R implementation of $\sim$ 2500 lines
- more examples
- regression tests