Power and sample size calculations for interaction analysis

In this package power and sample size calculations for the analysis of treatment-by-subgroup interactions are proposed. Therefore, the interaction effect is either defined via product-type interaction contrasts [1,2] or as ratio of treatment differences [3]. The latter formulation allows besides the detection of interactions also the assessment of the consistency/heterogeneity of the subgroup-specific treatment effect by the definition of an inconsistency margin. he methodology is applicable in trials with a continuous as well as a binary endpoint.

- calculate the power to detect an interaction effect
- calculate sample size to detect an interaction effect
- calculate power for consistency assessment
- calculate sample size for consistency assessment

Consistency assessment in multi-regional clinical trials

The function `PowCon` calculates the minimal power (any-pair power) associated with simultaneous tests for the ratios of treatment differences with a user defined inconsistency margin for the assessment of treatment-by-subgroup interactions.

```
#library(poco)
PowCon(mu=c(10,15,10,15,10,15,10,15,10,11),
       #vector of cell means from the cell
       #mu = means model corresponding to a two-way layout
       sd=sqrt(5),
       #number samples per treatment-by-subgroup combination
       n = 10,
       #pooled standard deviation
       n.sub=5,
       #number of levels of the subgrouping factor;
       #if n.sub=1 the argument SubMat is ignored
       TreatMat= "Tukey",
       #type of contrast for the treatment factor;
       #naming one of the contrast types available in contrMat(multcomp)
       SubMat = "GrandMean",
       #type of contrast for the subgrouping factor;
       #naming one of the contrast types available in contrMat(multcomp);
       #if n.sub=1 the argument SubMat is ignored
       thetas=1, #inconsistency margin(s)
       alpha=0.05, #familywise type I error to be controlled
       alternative="two.sided",#specifying the direction of the alternative hypothesis,
       #one of "two.sided","less","greater"
       type="anypair")#power definition, one of "global","anypair","allpair"
```

The function `nPowCon`calculates the sample size associated with simultaneous tests for the ratios of treatment differences with a user defined inconsistency margin for the assessment of treatment-by-subgroup interactions. The required size n is determined iteratively by starting with a given sample size and search until the power condition is satisfied.

```
nPowCon(min.power=0.9,#any pair power
       mu=c(10,15,10,15,10,15,10,15,10,11),#vector of cell means from the
       #cell means model corresponding to a two-way layout
       sd=sqrt(5),#pooled standard deviation
       n.sub=5,#number of levels of the subgrouping factor;
       #if n.sub=1 the argument SubMat is ignored
       TreatMat= "Tukey", #type of contrast for the treatment factor;
       #naming one of the contrast types available in contrMat(multcomp)
       SubMat = "GrandMean",#type of contrast for the subgrouping factor;
       #naming one of the contrast types available in contrMat(multcomp);
       #if n.sub=1 the argument
       thetas=1, #inconsistency margin(s)
       alpha=0.05, #familywise type I error to be controlled
       alternative="two.sided",##specifying the direction of the alternative hypothesis,
       #one of "two.sided","less","greater"
       type="anypair")#power definition, one of "global","anypair","allpair"
```

Qualitative interaction

Testing for or against a qualitative interaction is relevant in randomized clinical trials that use a common primary factor treatment and have a secondary factor, such as the centre, region, subgroup, gender or biomarker. Kitsche and Hothorn formulated interaction contrasts for ratios of differences between the levels of the primary treatment factor to detect a qualitative interaction. To calculate the power and sample size to detect a qualitative interaction the user has to set the inconsistency margin `thetas` to 1 and specify the `alternative` argument to `greater`. The approach is also applicable to detect treatment-by-subset interactions in a stratified, randomised clinical trial with a binary-response variable, see Kitsche (2014). The function `PowConBinom` calculates the any-pair power for multiple ratios of treatment differences for a binomial distributed endpoint. As illustrative example consider the Metoprolol Controlled-Release Randomized Intervention Trial in Heart Failure [5]. In this trial significant qualitative interactions were of particular interest, especially significant departures from the overall effect among any of the participating countries. The following code calculates the power to detect a qualitative interaction in this trial.

```
data(MetoCRXL2)#the data set is available from the poco package
library(MCPAN)#required package to estimate the success probabilities
MCPAN_Est <- binomest(Success ~ RegionTreat,data=MetoCRXL2, success="1", method="Wald")
MCPAN_Est$estp#estimated success probabilities
MCPAN_Est$n#sample size per treatment-by-region
PowConBinom(p=MCPAN_Est$estp,
            #binomial proportions for each tretament-by-subgroup combination
            n=MCPAN_Est$n, #sample sizes per treatment-by-subgroup combination
            n.sub = 12, #number of levels of the subgrouping factor;
            #if n.sub=1 the argument SubMat is ignored
            TreatMat = "Tukey", #type of contrast for the treatment factor;
            #naming one of the contrast types available in contrMat(multcomp)
            SubMat = "GrandMean",#type of contrast for the subgrouping factor;
            #naming one of the contrast types available in contrMat(multcomp);
            #if n.sub=1 the argument SubMat is ignored
            rhs = 0, #inconsistency margin(s)
            alternative = "less", #direction of the alternative hypothesis,
            #one of "two.sided","less","greater"
            alpha = 0.05,#familywise type I error to be controlled
            type="anypair")#power definition, one of "global","anypair","allpair"
```

Non-inferiority analysis

The presented methodology to assess the consistency of treatment effects is in general an extension of the problem of multiple testing for non-inferiority based on ratios was first adressed by Hauschke and Kieser. If the number of levels of the subgrouping factor is set to 1 the method simplifies to the approach of Dilba et al. [7]. The following code recalculates the method presented in Dilba et al..

```
PowCon(mu=c(10,9,9,11),
       sd=sqrt(5),
       n = 10,
       n.sub=1,
       TreatMat= "Dunnett",
       SubMat = "Tukey",
       thetas=0.8,
       alpha=0.05,
       alternative="greater",
       type="anypair")
```

References

1. Gabriel KR, Putter J,Wax Y. Simultaneous Confidence Intervals for product-type Interaction Contrasts. *Journal of the Royal Statistical Society Series B - Statistical Methodology* **1973**; 35(2):234-244.
2. Kitsche A, Schaarschmidt F. Analysis of Statistical Interactions in Factorial Experiments. *Journal of Agronomy and Crop Science* **2014**; early view
3. Kitsche A, Hothorn LA. Testing for qualitative interaction using ratios of treatment differences. *Statistics in Medicine* **2014**; 13(9):1477-1489
4. Kitsche A. Detecting qualitative interactions in clinical trials with binary responses. *Pharmaceutical statistics* **2014**; early view
5. MERIT-HF Study Group. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial In-congestive Heart Failure (MERIT-HF). *Lancet* **1999**; 353(9169):2001-2007.
6. Kitsche A, Power and sample size computations for simultaneous consistency assessment of treatment effects [manuscript in preparation]
7. Dilba G., Bretz F., Hothorn L.A., Guiard V. Power and sample size computations in simultaneous tests for non-inferiority based on relative margins. *Statistics in Medicine* **2006**; 25: 1131-1147