APPLYING ADDITIVE P-VALUES COMBINATIONS WITHIN STUDIES

GEORGETTE ASHERMAN (georgette@directeffects.net)
Direct Effects, LLC

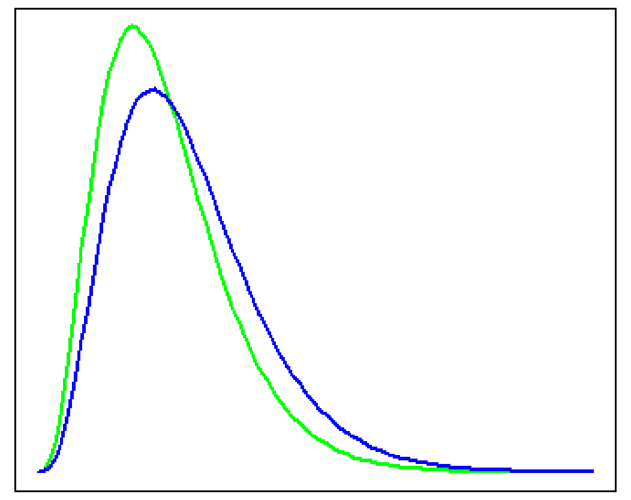
Direct Effects
Results-Driven Analytic & Data Services

Why this works

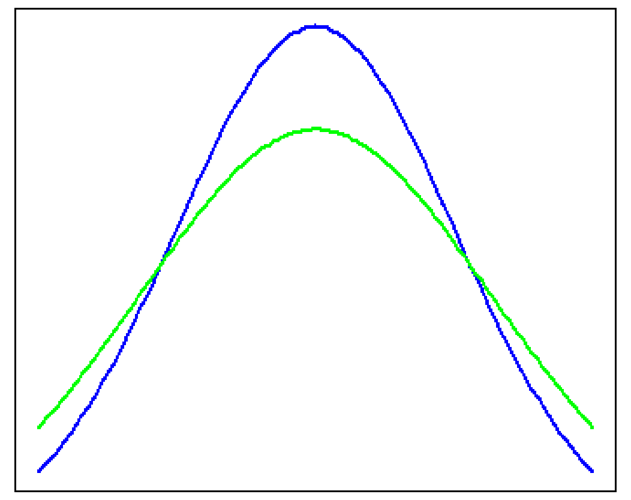
- P-values from a continuous test statistic ~ U(0,1) under the null hypothesis
- So the p-values can be transformed into the inverse of a cumulative distribution function $q_i = F^{-1}(p_i)$
- For k p-values, $C = \sum_{i=1}^{k} q_i$ and the overall test statistic is computed from the sampling distribution of C Examples are the Chi-square with 2k degrees of freedom and the Normal (0,k) distribution.

Adjusting for Correlated P-values

Chi-Square Density Graphs



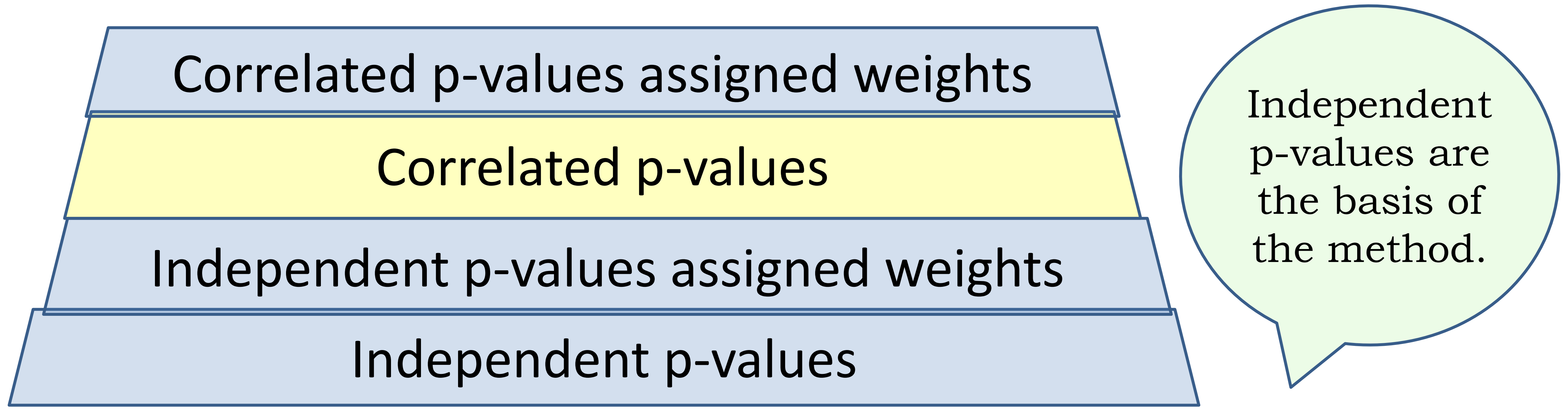
Normal Density Graphs



The green curves show an adjustment for correlation in the distribution of the sum. The X^2 changes parameters and the normal distribution widens.

What are Additive P-Value Combinations?

- P-value combinations perform a function on a vector of P-values to obtain a test of multiple hypothesis. These methods have a long legacy in statistics. They are common meta-analysis tools and have emerged in the 'omics area such as microarray tests.
- Additive P-value combinations add up transformed P-values to obtain a test statistic with a known distribution. The null hypothesis is that none of the tests are significant.
- There is an extensive literature for adjusting combined test statistics for correlation of P-values. Correlation matrices are obtained from formula or bootstrapping original analysis



Correlated p-values assigned weights

Correlated p-values

Independent p-values assigned weights

Independent p-values

Independent p-values are the basis of the method.

What advantages does this approach offer?

- Obtain an overall statistic when each test is distinct-from different studies or different modalities
- Avoids emphasis on any given test
- Retains original results-not changing P-value for each test
- Can be performed when raw data not available
- Can be done when modeling not feasible

Example with R

The demonstration is from a claim for a food additive on four biomarkers associated with a clinical outcome. The vector consists of 4 P-values from two-group, one-sided t-tests for a treatment effect with a baseline covariate (ANCOVA). Obtain upper tail P-value for uncorrelated P-values.

R-code

#assign vector, k =4

```
Pvec<-
c(0.1484,0.2807,0.0040,0.6657)
```

The Fisher Method- Chi Square (2k)

```
CFisher <- sum(-1*log(Pvec))
CombinedP1<-
pchisq(sum(CFisher, df=8,
lower.tail=FALSE)
```

CombinedP1 = 0.0195

The Stouffer Method- Normal (0,k)

```
CStouffer=sum(qnorm(Pvec,
   lower.tail=FALSE))
CombinedP2<-
pnorm(sum(CStouffer/sqrt(length
(Pvec)), lower.tail=FALSE)
```

CombinedP2 = 0.0263

Correlated p-value results on handout