The Use of Generalized P-Values and Generalized Confidence Intervals in Meta-Analysis

Guido Knapp, Bimal K. Sinha, Annette Böckenhoff

Department of Mathematics and Statistics,
University of Maryland, Baltimore County

Department of Statistics, University of Dortmund

GMDS 2006, Leipzig, September 10-14, 2006

Outline

Meta–Analysis

- combining results from k independent trials
- often published data:
  $\hat{\theta}_i$ estimate of treatment effect and
  $\hat{\sigma}_i^2$ estimate of variance of $\hat{\theta}_i$, $i = 1, \ldots, k$
- **common treatment effect** in all the trials: fixed effects model
  otherwise: random effects model, **overall treatment effect**,
  between–trial variance

Motivation

Hartung, Knapp (2001, Statist. Med.):
Effect measure: difference of normal means and assume a common treatment effect in all studies (fixed effects meta–analysis model)

- analysis in the random effects model is almost always better than analysis in the fixed effects model if the fixed effects model is the correct model
- use of an improved variance estimate of the overall treatment effect estimate often leads to more accurate results than use of the 'classical' variance estimate
- however, there is no 'clear' winner of the two random effects approaches in the true fixed effects model (conservative as well as anti–conservative results)

Motivation

Simulation results for six trials from Hartung, Knapp (2001):
Estimated confidence coefficients (in %) for $100\%(1-\alpha) = 95\%$

Pattern	ψ_{FE}	ψ_{RE}	ψ_q
1	87.57	93.63	93.68
2	86.53	92.90	92.75
3	91.94	94.93	94.17
4	91.78	94.58	94.45
5	94.66	96.08	94.95
6	94.42	96.15	94.87
7	92.01	94.53	93.97
8	90.77	94.53	94.71
9	92.75	95.19	94.35
10	93.95	95.76	94.56
11	92.72	95.08	94.83
12	94.45	96.14	94.96

Motivation

Idea:

- Does the use of generalized confidence intervals improve the analysis in the fixed effects model?

Exact Data Analysis

- generalized p–value (Tsui, Weehrahandi, 1989, JASA)
- generalized confidence interval (Weerahandi, 1993, JASA)
- Weerahandi (1995): Exact Statistical Methods for Data Analysis. Springer:New York
- application in biometry:
  - exact inference for growth curves (Weerahandi, Berger, 1999, Biometrics)
  - interval estimation and hypothesis testing of intraclass correlation coefficient (Tian, Cappelleri, 2004, Statist. Med.)

Definition

Suppose that $X = (X_1, X_2, \ldots, X_n)$ forms a random sample from a distribution which depends on the parameters $\psi = (\theta, \nu^T)^T$ where θ is the parameter of interest and ν^T is a vector of nuisance parameters.

A generalized pivot $R(X; x, \theta, \nu)$, where x is an observed value of X, has the following two properties:

1. $R(X; x, \theta, \nu)$ has a distribution free of unknown parameters.
2. The value of $R(x; x, \theta, \nu)$ is θ.

Definition

A generalized pivot $R(X; x, \theta, \nu)$, where x is an observed value of X, has the following two properties:

1. $R(X; x, \theta, \nu)$ has a distribution free of unknown parameters.
2. The value of $R(x; x, \theta, \nu)$ is θ.

Let R_α be the (100α)th percentile of R.
Then $(R_{\alpha/2}, R_{1-\alpha/2})$ becomes a $100\%(1-\alpha)$ two–sided generalized confidence interval for θ.

Random effects model

Let us consider k independent trials.
For $i = 1, \ldots, k$,

$$\hat{\theta}_i \sim \mathcal{N}\left(\theta_i\,,\,\sigma_i^2\right)$$

and

$$\theta_i \sim \mathcal{N}\left(\theta\,,\,\tau^2\right)$$

θ	—	overall treatment effect, parameter of interest
τ^2	—	heterogeneity parameter
σ_i^2	—	within study variance, $i = 1, \ldots, k$

Random effects model

For $i = 1, \dots, k$,

$$\hat{\theta}_i \sim \mathcal{N}\left(\theta\,,\ \tau^2 + \sigma_i^2\right)$$

Parameters $\psi = (\theta, \nu)$, $\nu^T = (\tau^2, \sigma_1^2, \dots, \sigma_k^2)$

Generalized Pivot

For $i = 1, \ldots, k$, let $\bar{X}_i$ and $\bar{Y}_i$ be sample means, S^2_{Xi} and S^2_{Yi} be the sample variances, σ^2_{Xi} and σ^2_{Yi} be the population variances, and n_{Xi} and n_{Yi} be the sample sizes for treatment group and control group, respectively.

Let θ be the difference of means, then

$$D_i = \bar{X}_i - \bar{Y}_i \sim \mathcal{N}\left(\theta, \frac{\sigma^2_{Xi}}{n_{Xi}} + \frac{\sigma^2_{Yi}}{n_{Yi}} + \tau^2\right)$$

Generalized Pivot

Suppose $\tau^2 = 0$, that is,

$$D_i = \bar{X}_i - \bar{Y}_i \sim \mathcal{N}\left(\theta, \frac{\sigma^2_{Xi}}{n_{Xi}} + \frac{\sigma^2_{Yi}}{n_{Yi}}\right)$$

Then,

$$D_w = \sum_{i=1}^{k} \frac{w_i D_i}{\sum_j w_j}, \quad w_i = \left(\frac{\sigma^2_{Xi}}{n_{Xi}} + \frac{\sigma^2_{Yi}}{n_{Yi}}\right)^{-1}$$

and

$$Z = \frac{D_w - \theta}{1/\sqrt{\sum_j w_j}} \sim \mathcal{N}(0,1)$$

Generalized Pivot

$$V_{Xi} = (n_{Xi} - 1) S_{Xi}^2 / \sigma_{Xi}^2 \sim \chi^2_{n_{Xi}-1}$$

$$V_{Yi} = (n_{Yi} - 1) S_{Yi}^2 / \sigma_{Yi}^2 \sim \chi^2_{n_{Yi}-1}$$

generalized pivots:
let s_{Xi}^2 and s_{Yi}^2 denote the observed values for S_{Xi}^2 and S_{Yi}^2

$$R_{\sigma^2_{Xi}} = \frac{(n_{Xi} - 1) s_{Xi}^2}{V_{Xi}} \sim \frac{(n_{Xi} - 1) s_{Xi}^2}{\chi^2_{n_{Xi}-1}}$$

$$R_{\sigma^2_{Yi}} = \frac{(n_{Yi} - 1) s_{Yi}^2}{V_{Yi}} \sim \frac{(n_{Yi} - 1) s_{Yi}^2}{\chi^2_{n_{Yi}-1}}$$

Generalized Pivot

Define R_{w_i}

$$R_{w_i} = 1/(R_{\sigma^2_{Xi}}/n_{Xi} + R_{\sigma^2_{Yi}}/n_{Yi})$$

and, with d_i the observed value of D_i,

$$d_{R_w} = \sum_{i=1}^{k} \frac{R_{w_i} d_i}{\sum_j R_{w_j}}$$

Generalized pivotal quantity

$$R_\theta^Z = d_{R_w} - \frac{Z}{\sqrt{\sum_j R_{w_i}}}$$

Generalized pivot

$$R_\theta^Z = d_{R_w} - \frac{Z}{\sqrt{\sum_j R_{w_i}}}$$

- the distribution of R_θ^Z is independent of any unknown parameters
- the value of R_θ^Z is θ as $D_i = d_i$, $S_{Xi}^2 = s_{Xi}^2$, and $S_{Yi}^2 = s_{Yi}^2$, $i = 1, \ldots, k$.

Computing algorithm

Given data $(\hat{\theta}_i, \hat{\sigma}^2_{Xi}, \hat{\sigma}^2_{Yi}, n_{Xi}, n_{Yi})$:

1. For $i = 1, \ldots, k$, generate $V_{Xi} \sim \chi^2_{n_{Xi}-1}$, $V_{Yi} \sim \chi^2_{n_{Yi}-1}$. Compute $R_{\sigma^2_{Xi}}$ and $R_{\sigma^2_{Yi}}$
2. Calculate R_{w_i} for $i = 1, \ldots, k$, and d_{R_w}.
3. Generate $Z \sim \mathcal{N}(0, 1)$. Compute R^Z_θ.
4. Repeat step 1-3 a total of m times
5. Rank the array of R^Z_θ.
6. Compute the percentiles $(R^Z_\theta(\alpha/2), R^Z_\theta(1 - \alpha/2))$.

Simulation study

- fixed effects model with $k = 6$ trials
- effect measure: difference of normal means $\theta = 0$
- calculation of a single generalized confidence interval based on $m = 5.000$ replications
- all estimated confidence coefficients based on 10.000 simulations runs
- various combinations of sample sizes and within–trial variances

Results (I)

- balanced sample sizes $n_{Xi} = n_{Yi} = n$, $i = 1, \ldots, k$
- homoscedastic variances $\sigma^2_{Xi} = \sigma^2_{Yi} = 1$, $i = 1, \ldots, k$

Estimated confidence coefficients (in %) given a nominal confidence coefficient of 95% and average length (in parentheses)

Sample size	ψ_{FE}	ψ_{RE}	ψ_q	gen CI
5	86.62	92.88	93.09	95.63
	(0.90)	(1.07)	(1.23)	(1.17)
10	92.05	94.96	94.30	95.26
	(0.68)	(0.77)	(0.88)	(0.76)
20	93.93	95.68	94.71	95.21
	(0.49)	(0.55)	(0.62)	(0.52)
40	93.84	95.60	94.60	94.66
	(0.35)	(0.39)	(0.45)	(0.36)

Results (II)

- unbalanced sample sizes $n_{Xi} = n_{Yi} = n_i$, $i = 1, \ldots, k$
- heteroscedastic variances $\sigma^2_{Xi} = \sigma^2_{Yi} = \sigma^2_i$, $i = 1, \ldots, k$

Estimated confidence coefficients (in %) given a nominal confidence coefficient of 95% and average length (in parentheses)

Pattern	ψ_{FE}	ψ_{RE}	ψ_q	gen CI
1	89.68	94.05	93.86	95.67
	(0.93)	(1.08)	(1.23)	(1.12)
2	90.65	94.66	94.03	95.41
	(0.85)	(1.01)	(1.19)	(0.99)
3	90.98	94.54	93.98	94.74
	(0.76)	(0.89)	(1.02)	(0.86)
4	93.18	95.79	95.25	94.67
	(0.63)	(0.87)	(1.09)	(0.67)

Main Conclusions

- Generalized confidence intervals for the difference of normal means is an efficient procedure in the fixed effects meta–analysis model
- Procedure for the difference of normal means is based on exact distributions

Outlook

- Extension to random effects model: adapt the proposal in Iyer et al. (2004, JASA)
- Other effect measures:
  - no exact normal and χ^2 distribution, approximation needed
  - investigation of the performance is still to be done

Main References

- Hartung, J., Knapp, G. (2001). On tests of the overall treatment effect in the meta-analysis with normally distributed responses. *Statistics in Medicine* **20**, 1771-1782.
- Iyer, H., Wand, J., Mathew, T. (2004). Models and confidence intervals for true values in interlaboratory trials. *Journal of the American Statistical Association* **99**, 1060-1071.
- Tsui, K., Weerahandi, S. (1989). Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association* **84**, 602-607.
- Weerahandi, S. (1993). Generalized confidence intervals. *Journal of the American Statistical Association* **88**, 899-905.