A parametric Rüger test and diagnostic trials

Robin Ristl

Section for Medical Statistics, Center of Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna

joint work with
Christian Gartner, Thomas Lang and Martin Posch

ROeS 2013

Motivation: Trial of a medical diagnostic tool

Trial plan:

- Apply diagnostic tool to 200 patients
- Assess output by $n = 3$ readers

Each patient i is classified by each reader j:

- $y_{ij} = 1$ diseased
- $y_{ij} = 0$ healthy or

Measure of performance

- $Sensitivity_j = P(Y_{ij} = 1 | diseased)$
- Estimate from m diseased patients:
  $\hat{s}_j = \frac{1}{m} \sum_{i=1}^{m} y_{ij}$

The proposed testing procedure

Aim: Show that a sensitivity > 0.7 can be achieved

- Hypothesis for reader j $H_j : sens_j \leq 0.7$
- Global null hypothesis $H : H_1 \cap H_2 \cap H_3$
- Test each H_j seperately at level α
- If the test is significant for 2 out of 3 readers, reject global null hypothesis

Question:

What is the type I error rate of this procedure?

It is a Rüger testing procedure

- Perform a level α test for each H_j
- Reject $H : H_1 \cap \ldots \cap H_n$ if k out of n individual hypotheses H_j are rejected at level α
- A sharp upper bound for the global level is $\alpha \frac{n}{k}$ (Rüger 1978)
- Test individual hypothesis on level $\alpha \frac{k}{n}$ to maintain a global level α
- k must be prespecified
- Special case $k = 1$ gives Bonferroni test

Answer, without distributional assumptions:

Type I error rate $\leq \alpha \frac{3}{2}$

Estimates of performance asymptotical multivariate normal

Sensitivity for reader j, m diseased patients:

Estimate: $\hat{s}_j = \frac{1}{m}\sum_{i=1}^{m} y_{ij}$
Null hypothesis: $H_j : s_j = s_0$
Test statistic $x_j = (\hat{s}_j - s_0)/SE(\hat{s}_j)$

Asymptotically, by multivariate c.l.t.

$$(X|H) \sim N_3(0, \Sigma)$$

Aim:

See if we can refine boundary of type I error rate under assumption of normality

Reject H if 2 out of 3 H_j are significant

Reject H_j if $x_j > c = \Phi^{-1}(1-\alpha)$

Probability to reject $H : H_1 \cap H_2 \cap H_3$

$$\pi = P(X_1 > c, X_2 > c) + P(X_1 > c, X_3 > c) + P(X_2 > c, X_3 > c) - 2P(X_1 > c, X_2 > c, X_3 > c)$$

Simple special cases:

X uncorrelated, $\alpha \in [0; 0.5]$: $\pi = 3\alpha^2 - 2\alpha^3 < \alpha$

Perfect correlation: $\pi = \alpha$, since all X_j identical

Perfect correlation of any pair $(X_s, X_t), s \neq t$: $\pi = \alpha$

Equally correlated test statistics

Assumptions:

$\alpha \in [0; 0.5]$

Test statistics $(X|H) \sim N_3(0, \Sigma)$

Equal correlations $cor(X_i, X_j) = \rho, i \neq j$

Proposition 1

Under the global null hypothesis and the above assumptions, the probability π to reject at least two out of three one-sided hypotheses at a level α is equal or below α.

Lemma for Derivative of π

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N\left(\mu = 0, \Sigma = \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix}\right), \alpha \in [0; 0.5]$$

then

$$\frac{\partial \pi}{\partial \rho_i} \geq 0$$

if

$$\rho_j + \rho_k \leq 1 + \rho_i$$

for (i, j, k) being any permutation of $(1, 2, 3)$

- Holds for $\rho_i = \rho_k = \rho_j \leq 1$
- So π is monotonously increasing in ρ
- $\pi = \alpha$ for $\rho = 1$
- Therefore $\pi \leq \alpha$, proof of proposition 1

Location of extrema in general case

Proposition 2

Under the global null hypothesis, all local extreme values of π in (ρ_1, ρ_2, ρ_3) are on the boundary of the parameter space.

Proof: A local extreme value not on the boundary means $\frac{\partial \pi}{\partial \rho_i} = 0, i = 1, 2, 3$. Using Lemma this means

$$\rho_1 + \rho_2 = 1 + \rho_3$$
$$\rho_1 + \rho_3 = 1 + \rho_2$$
$$\rho_3 + \rho_2 = 1 + \rho_1$$

Unique solution $(\rho_1, \rho_2, \rho_3) = (1, 1, 1)$ on boundary.

Conclusions form theoretical part

- In the case of multivariate normal statistics with equal correlations, using the 2 out of 3 Rüger test with individual level α is conservative at global level α.
- For general correlation structures, the maximum of the probability to falsely reject H is found on the boundary of the correlation parameter space.
- General case is studied by considering the boundary.

Numeric solution for general correlation structure

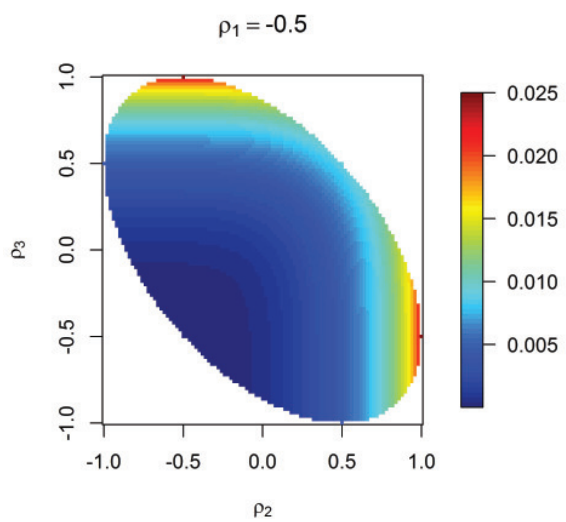
- Calculate probability to falsely reject H numerically (R package mvtnorm and TVPACK algorithm)
- One-sided level $\alpha = 0.025$
- Correlation matrices $\Sigma = \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix}$, n.n.d.

  with
  $\rho_{1,2,3} \in \{-1, -0.98, -0.96, \ldots 0.96,\ 0.98,\ 1\}$

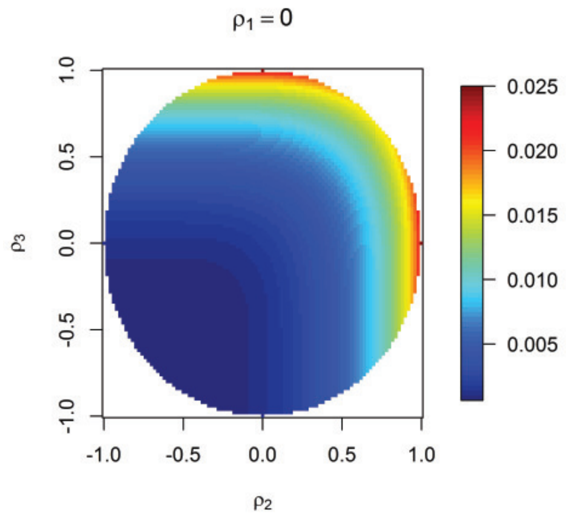
Numeric answer

$max(\pi) = 0.025$, found for any $\rho_i = 1$

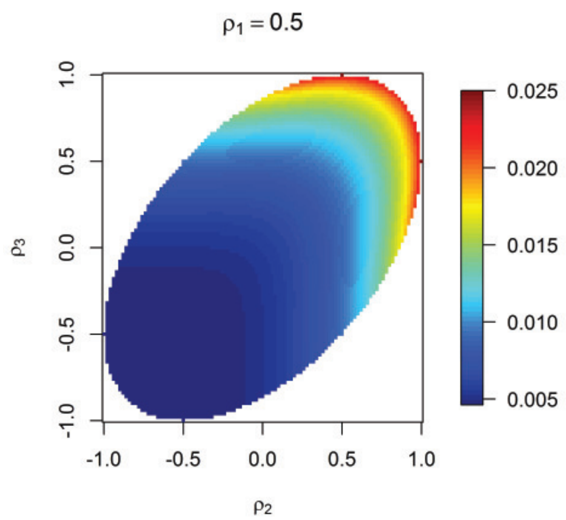
Probability to reject H - Numeric results for $\alpha = 0.025$ (1)



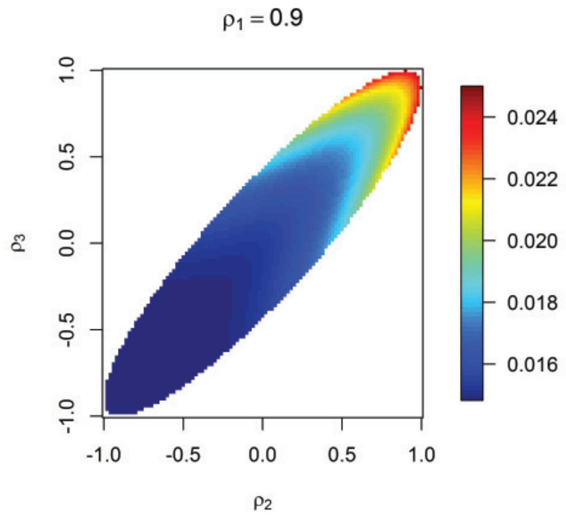
Probability to reject H - Numeric results for $\alpha = 0.025$ (2)



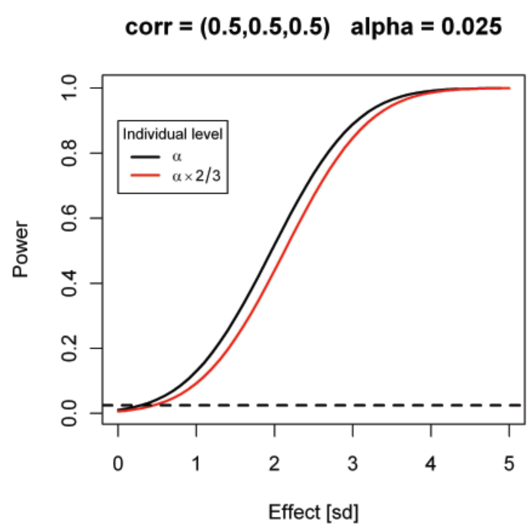
Probability to reject H - Numeric results for $\alpha = 0.025$ (3)



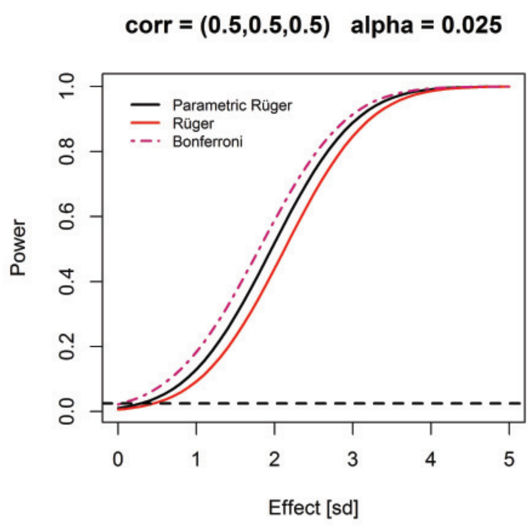
Probability to reject H - Numeric results for $\alpha = 0.025$ (4)



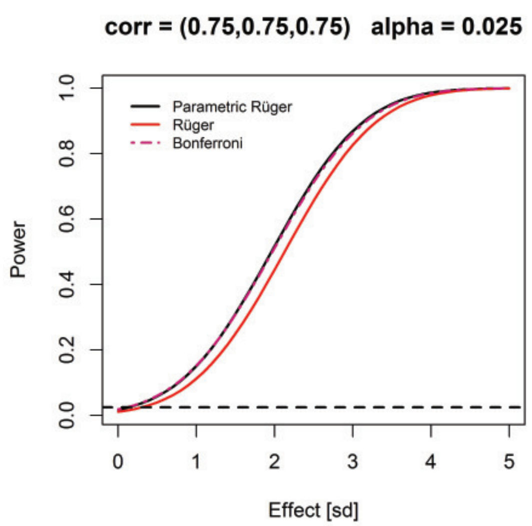
Power, equal correlations $\rho = 0.5$, all H_i false



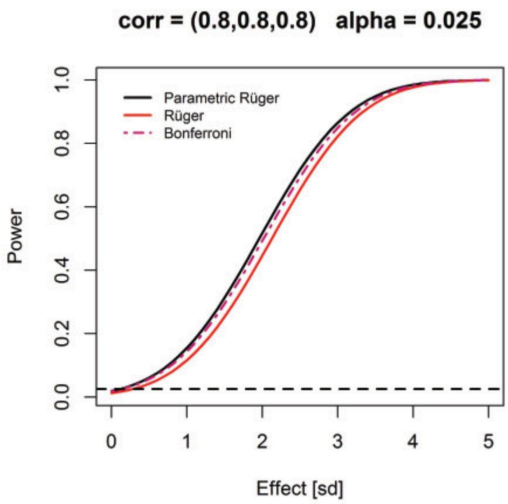
Comparing with Bonferroni test (1)



Comparing with Bonferroni test (2)



Comparing with Bonferroni test (3)



Parametric Rüger test is superior to Bonferroni for high correlations and common effect only.

Note: Bonferroni reaches maximal type I error rate very close to α for $\rho_1 = \rho_2 = \rho_3 = -0.5$

Discussion

- Several readers typical for diagnostic trial
- Specific scenario for three readers
- Rüger test free of distributional assumptions:
  Type I error rate$\leq \alpha\frac{3}{2}$
- May assume multivariate normality
- Result for this parametric Rüger test:
  Type I error rate$\leq \alpha$
- Parametric test has higher power

Literature

Caraux, G. and Gascual, O. (1992), “Bounds on distribution functions of order statistics for dependent variates”, Statistics & Probability Letters 14, 103-105

Committee for medicinal products for human use, European Medicines Agency (2009), “Appendix 1 to the guideline on clinical evaluation of diagnostic agents (CPMP/EWP/1119/98 REV. 1) on imaging agents,” Doc. Ref. EMEA/CHMP/EWP/321180/2008.

Genz, A. (2004), “Numerical computation of rectangular bivariate and trivariate normal and t probabilities”, *Statistics and Computing*, 14, 251-260.

Hommel, G., Bretz, F. and Maurer W. (2011), “Multiple hypotheses testing based on ordered p values—A historical survey with applications to medical research”, *Journal of Biopharmaceutical Statistics*, 21, 595-609.

Plackett, R. L. (1954), “A reduction formula for normal multivariate integrals”, *Biometrika*, 41, 351-360.

Rüger, B. (1978), ”Das maximale Signifikanzniveau des Tests: ,,Lehne H0 ab, wenn k unter n gegebenen Tests zur Ablehnung führen”, *Metrika*, 25, 171-178.

Thank you for your attention.

Outline for proof of Lemma

Plackett (1954) has shown $\frac{\partial \phi}{\partial cov(x_i, x_j)} = \frac{\partial^2 \phi}{\partial x_i \partial x_j}$

Applying to p gives, e.g.

$$\frac{\partial p}{\partial \rho_1} = \frac{\partial^2 p}{\partial x_1 \partial x_2} =$$

$$\int_{-\infty}^{\infty} \phi(z, z, x_3)\, dx_3 - 2 \int_{z}^{\infty} \phi(z, z, x_3)\, dx_3$$

$$\frac{\partial p}{\partial \rho_1} > 0 \text{ if } z > E(x_3 | x_1 = z, x_2 = z)$$

For $z > 0$, this is equivalent to $\rho_j + \rho_k < 1 + \rho_i$