Which Multiple Testing Methods are Optimal?

Peter H. Westfall,
Texas Tech University

Background

- The scientific literature has recently experienced an embarrassment of contradictory results:

- Ioannidis, J.P. (2005), "Contradicted and Initially Stronger Effects in Highly Cited Clinical Research,"*J. Amer. Med. Assoc.* 294, 218--228.
- Bertram, L., McQueen, M. B., Mullin, K., Blacker, D., and Tanzi, R. E. (2007), "Systematic Meta-analyses of Alzheimer Disease Genetic Association Studies: the AlzGene Database," *Nature Genetics* 39, 17--23.
- Boffetta, P., McLaughlin, J.K., La Vecchia, C., Tarone, R.E., Lipworth, L., Blot, W. J., (2008), "False-Positive Results in Cancer Epidemiology: A Plea for **Epistemological Modesty**," *J. Nat. Cancer Inst.* 100, 988--995.

Goals

- Compare fixed critical value methods in terms of loss

Q: Does *m* matter? Do **data** correlations matter?

A: It depends on how you feel about type I versus type II errors (i.e., relative costs)

Background

- “Lehmann (1957a,b) was the first to consider multiple comparisons from a decision-theoretic viewpoint.”

  – Hochberg and Tamhane (1987), *Multiple Comparisons Procedures* (Wiley)

Data Setup of this Talk

Data:

$\mathbf{z} \mid \boldsymbol{\theta} \sim N_m(\boldsymbol{\theta}, \rho)$, ρ a correlation matrix.

Model:

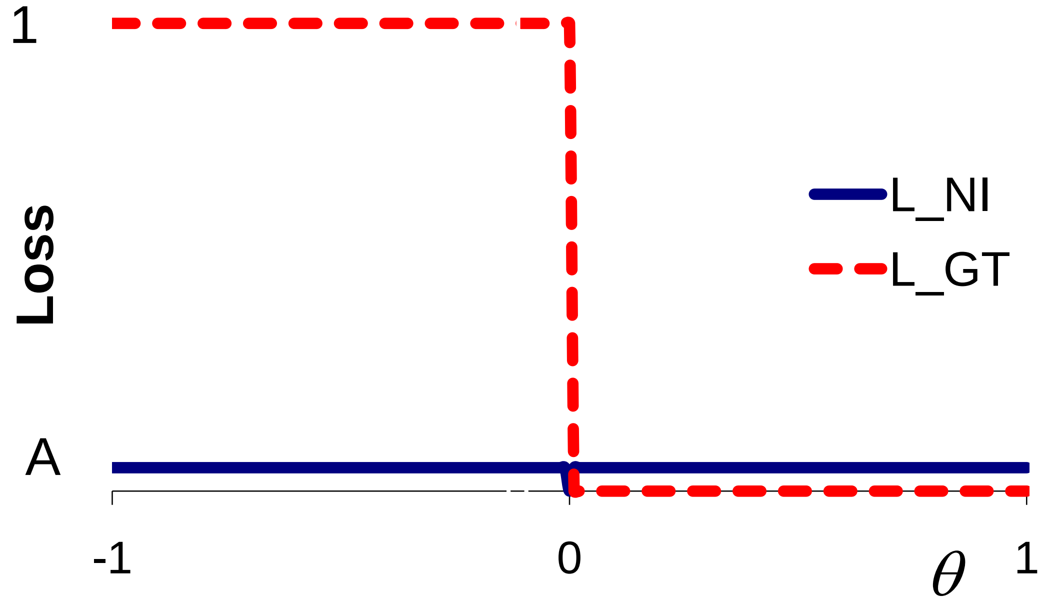
$\theta_i \sim_{\text{iid}} N(0, \sigma^2)$, σ^2 known.

Decision Theory

- Lehmann (1957a,b) Annals
- Hochberg and Tamhane (1987)
- Three-decision problem: Decide either
  - GT: $\theta_i > 0$
  - LT: $\theta_i < 0$, or
  - NI: $\theta_i \sim 0$ (or “EM”)

A Component Loss Function

- $L_{\mathrm{GT}}(\theta)$, $L_{\mathrm{LT}}(\theta)$, $L_{\mathrm{NI}}(\theta)$; for example:



Actual and Expected Loss

- Actual loss using method "M":

$$L_i^{(\mathrm{M})}(\mathbf{z},\theta_i)$$
$$= I_i(GT\,|\,\mathbf{z})L_{GT}(\theta_i)$$
$$+I_i(LT\,|\,\mathbf{z})L_{LT}(\theta_i)$$
$$+I_i(NI\,|\,\mathbf{z})L_{NI}(\theta_i)$$

- Expected Loss: $\Psi_i^{(\mathrm{M})} = E_{\mathbf{z},\theta_i}\left(L_i^{(\mathrm{M})}(\mathbf{z},\theta_i)\right)$
- Combined Loss: $\Psi^{(\mathrm{M})} = \sum \Psi_i^{(\mathrm{M})}$ (additive!?)

Decision Rules

- Decide
  - LT if $z_i < -c$
  - GT if $z_i > c$
  - NI if $-c \le z_i \le c$
- If $\boldsymbol{\rho} = \mathbf{I}$, then $c = (1 + 1/\sigma^2)z_{1\text{-}A}$ is optimal.

$\Rightarrow$ For Bonferroni-like procedures to be optimal, A=A(m).

Does *m* Matter?

- Theorem: If $A(m) = o(1)$ and $1/A(m) = o(m\{\ln(m)\}^{1/2})$, then $\Psi^{(Bon)} \sim \Psi^{(Optimal)}$.

  $\Rightarrow$ If the loss of a single Type I error equals βm Type II errors ($0<\beta<1$), then Bonferroni is optimal and fixed significance level procedures (like FDR) are inadmissible.

Lu, Y., and Westfall, P. (2009). Is Bonferroni Admissible for Large *m*? *American Journal of Mathematical and Management Sciences*, Vol. 29 (1&2), 51-69.

From Lu, Y., and Westfall, P. (2009). Is Bonferroni Admissible for Large *m*?

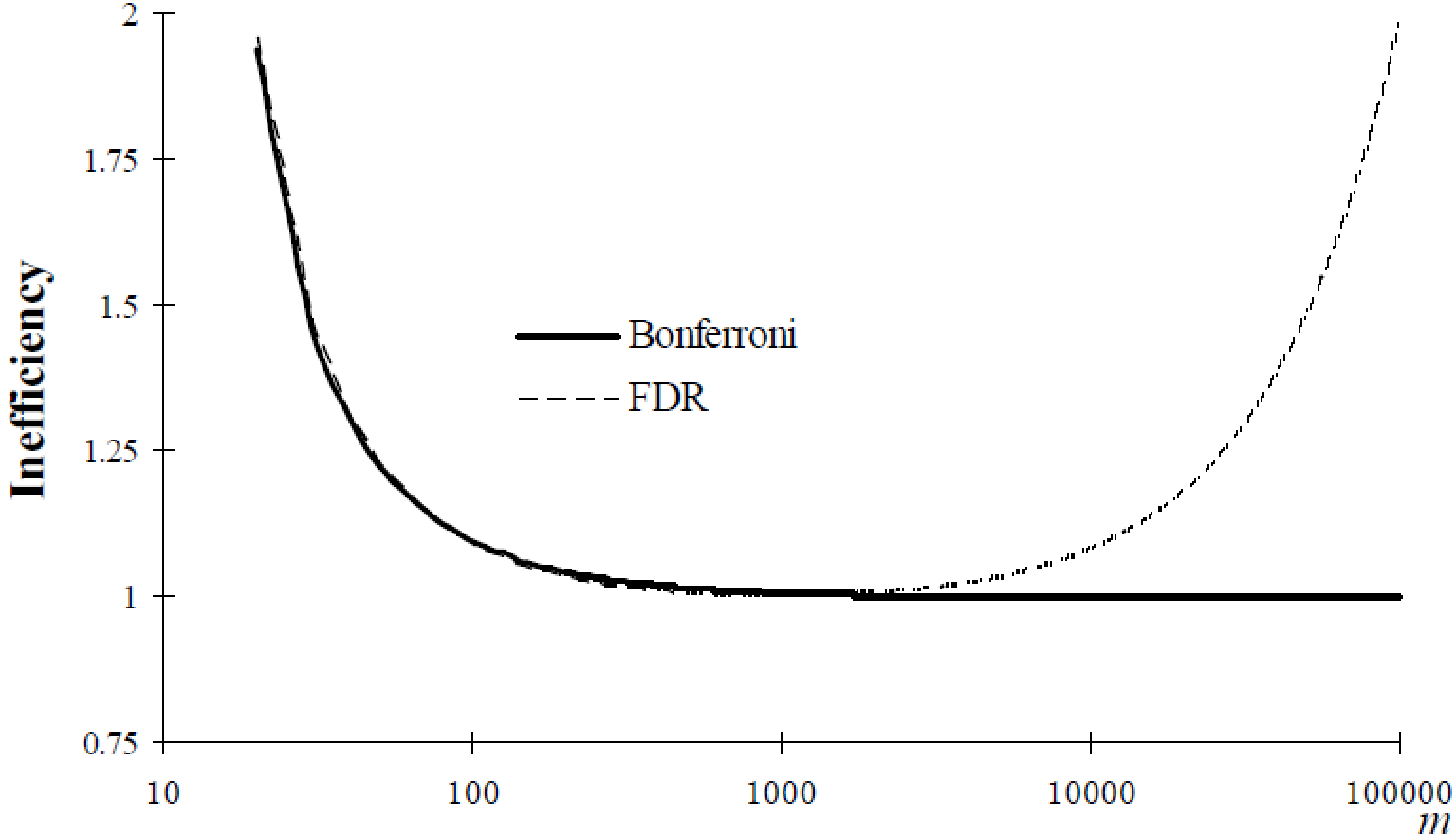


FIGURE 4. Ratio of Risk Relative to Optimal Rule

Do Data Correlations Matter?

“Reject H_i” if $|z_i|>c$, $i=1,\ldots,m$.

Let V = number of false discoveries.

With higher correlations among z’s:

- E(V) is unaffected
- P(V>0) is lower (smaller FWER)
- Var(V) is higher (potentially high # of false discoveries)

Effect of Correlation with Additive Loss

- No affect on expected value $\Rightarrow$ optimal *c* not affected

- Affects percentiles $\Rightarrow$ optimal *c* *is* affected

VaR = “Value at risk”=95th pctle of Loss (finance)

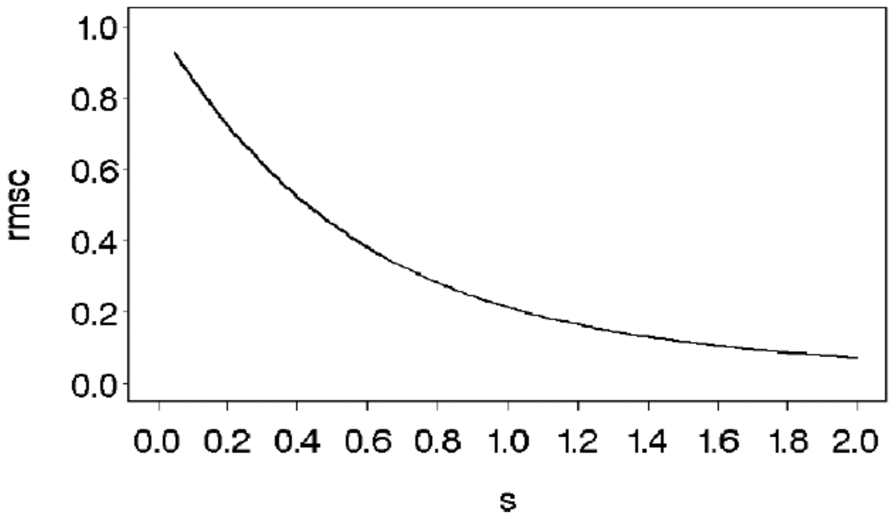
A Model for Studying Effect of Correlation

Suppose $\mathbf{z} \mid \boldsymbol{\theta} \sim N_m(\boldsymbol{\theta}, \rho)$, with $\rho = \lambda\lambda' + \psi^2$, λ (m x1) and ψ^2 diagonal.

Then $\rho_{ij} = \lambda_i \lambda_j$.

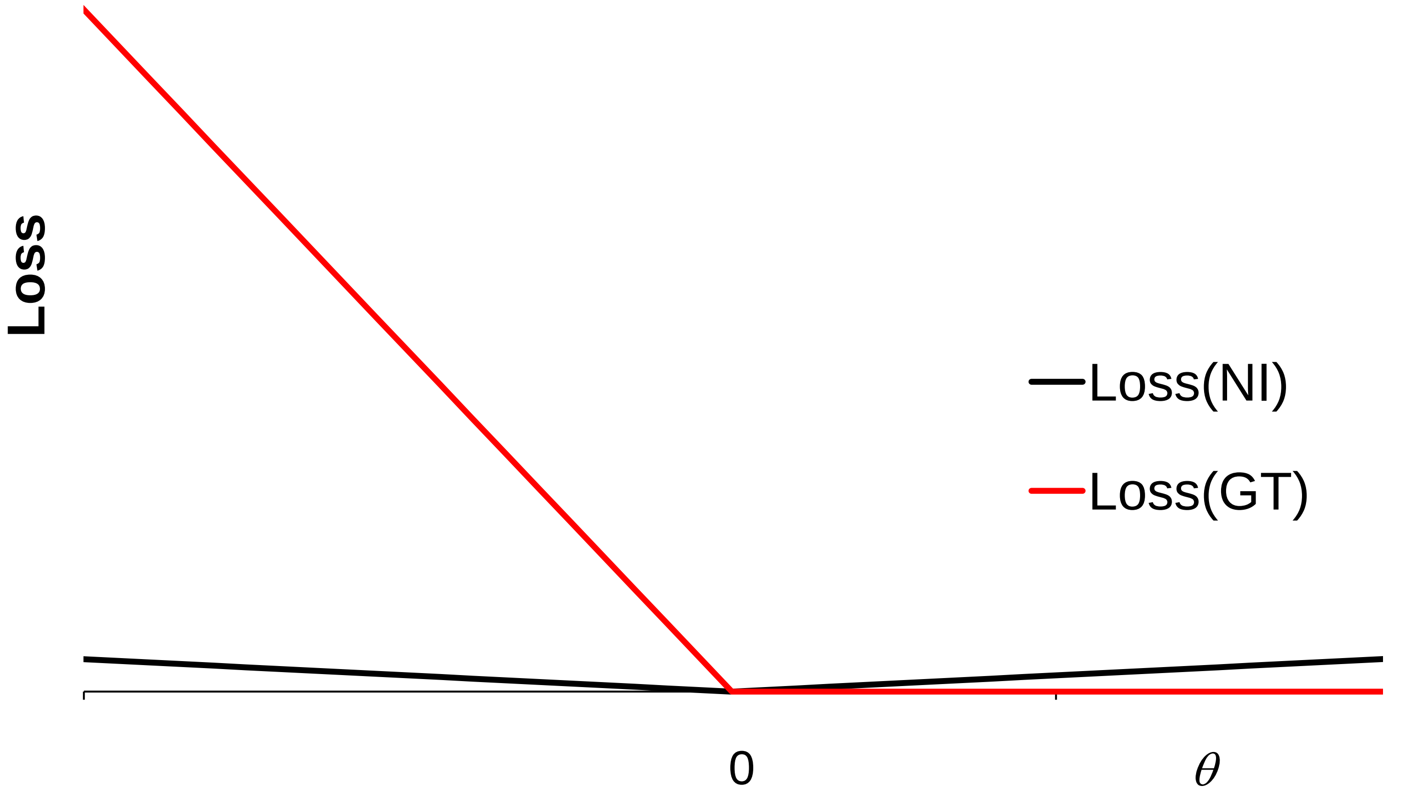
Let $\lambda_i = U_i / (U_i^2 + s^2)^{1/2}$, where $U_i \sim_{\text{iid}} U(-1,1)$.

Then $E(\rho_{ij})=0$ and $\text{rmsc} \equiv \left\{ E(\rho_{ij}^2) \right\}^{1/2} = 1 - s \tan^{-1}(1/s)$.

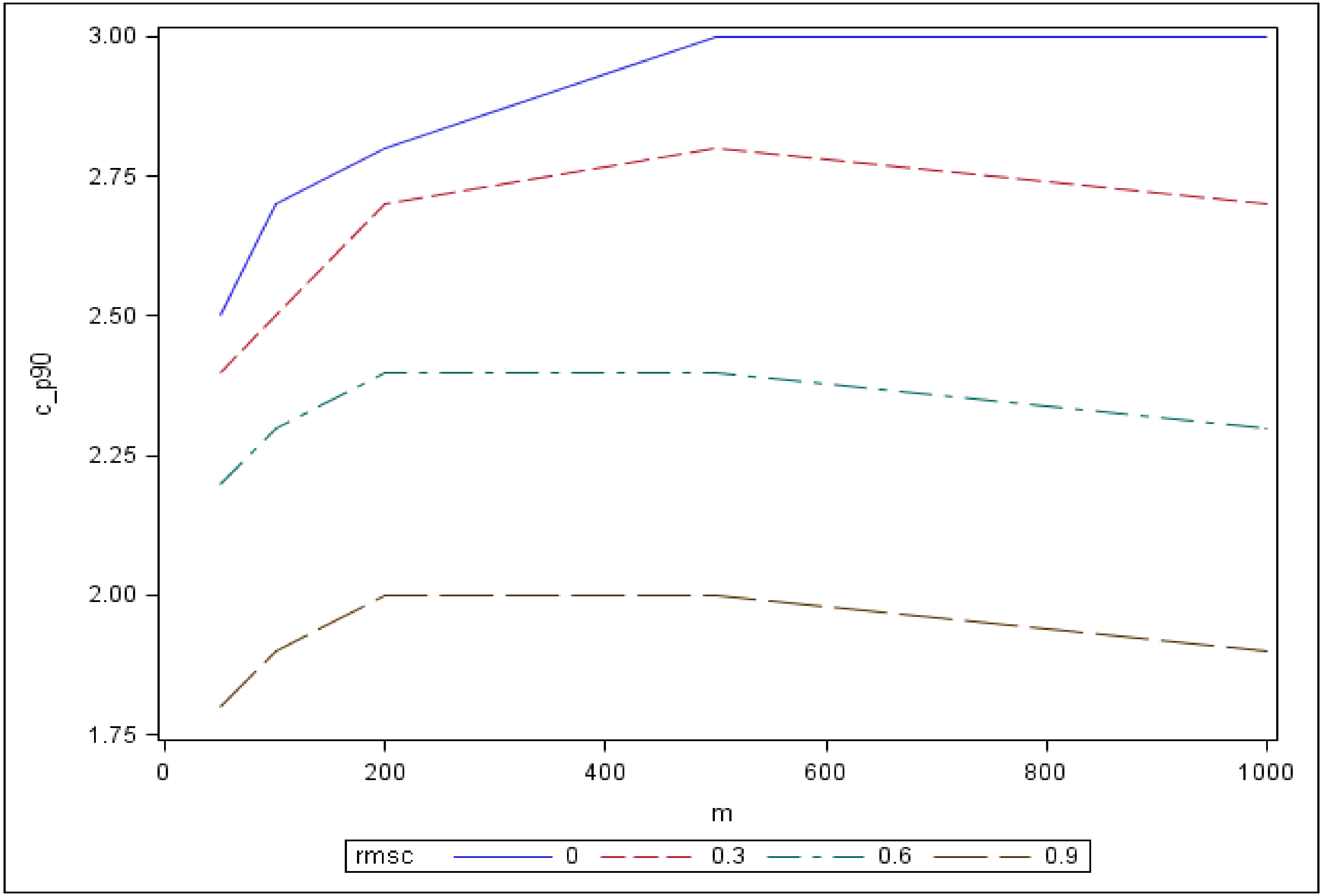


Waller-Duncan Loss

$L_{GT}(\theta) = -(K+1)\theta, \quad \theta < 0; \;\; L_{GT}(\theta) = 0, \;\; \text{o/w.}$
$L_{NI}(\theta) \; = \; |\theta| \, .$



90^{th} Pctle-Minimizing Optimal *c*, K=100



Should Loss Be Additive?

- Is the cost difference between 10 and 11 Extraterrestrial Intelligence claims the same as the cost difference between 0 and 1?
- Is the cost difference between 10 and 11 shouts of “fire” in a crowded theater the same as the cost difference between 0 and 1?

‘Fire-In-The-Theater’ Loss Function

Let n_1 = # Directional Errors

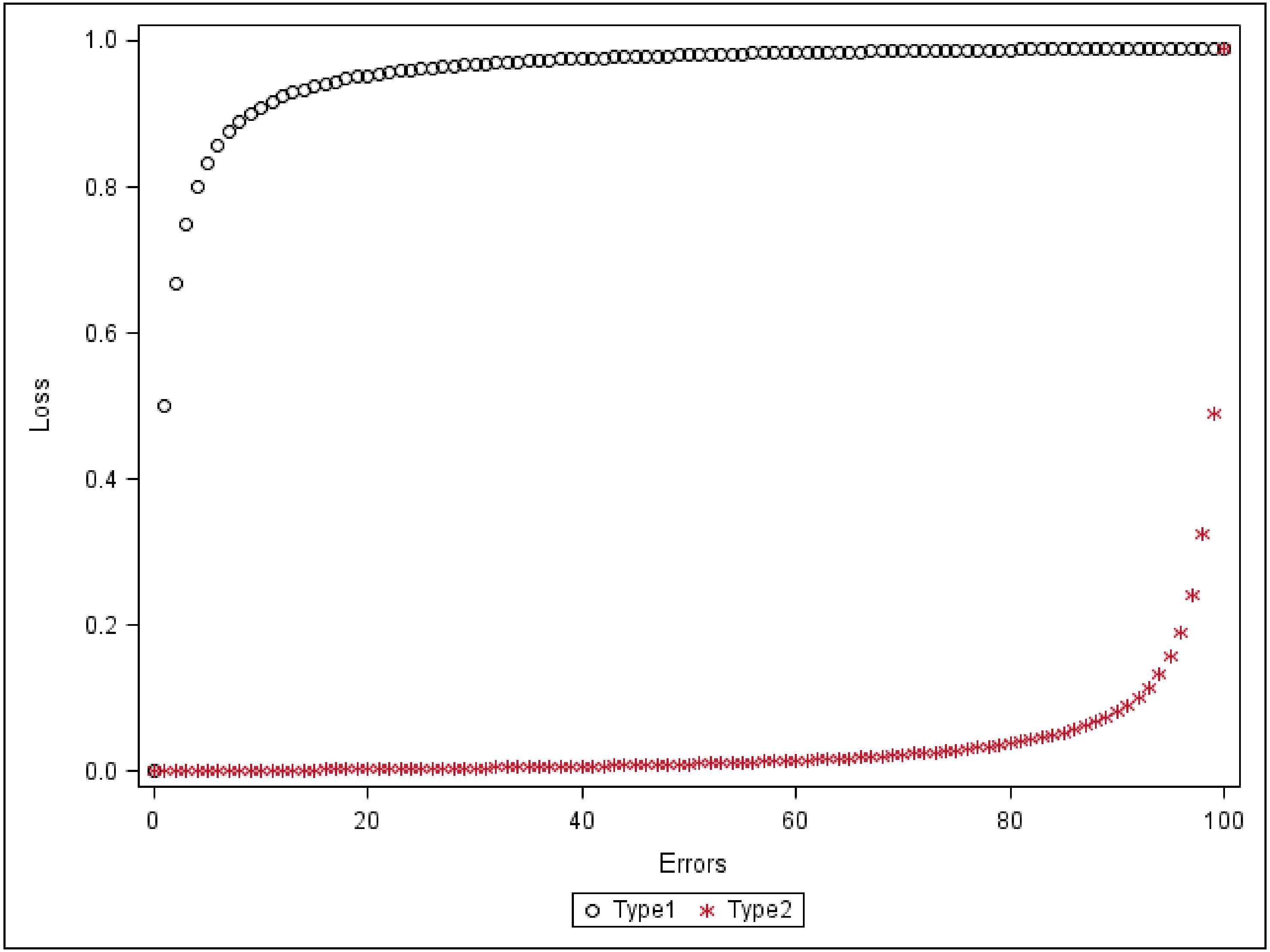
Let n_2 = # “Not Interesting” claims

$L_1 = n_1/(n_1 + 1)$

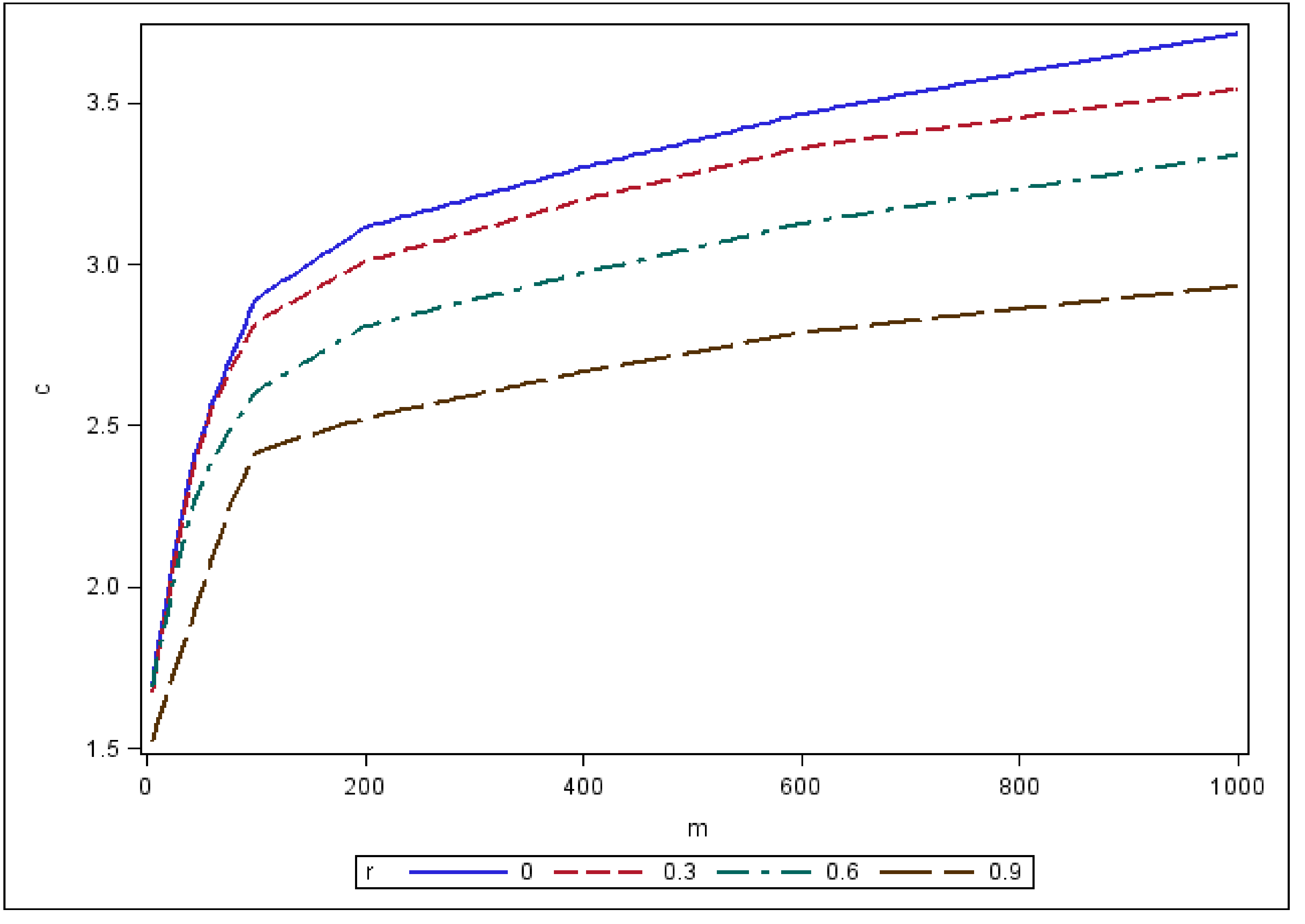
$L_2 = 1/(m - n_2 + 1) - 1/(m + 1)$

“Fire in the Theater” Loss = $L_1 + L_2$

Fire-In-The-Theater Loss Function Components, *m*=100



Expected Value-Minimizing Optimal *c* for Fire-In-The-Theater Loss Function



Conclusions

If Type I errors are serious then:

1. m matters: larger c needed with larger m.

2. Data correlation matters: smaller c allowed with higher data correlation.