# On the Optimum Number of Hypotheses to Test when the Number of Observations is Limited

A. Futschik and M. Posch

Vienna University & Medical Univ. of Vienna

# A Main Goal in Statistics

- Extract as much information as possible from a limited number of observations
- In the context of Multiple Hypothesis Testing: Reject (correctly!) as many null hypotheses as possible while still ensuring some global control of the type I error.

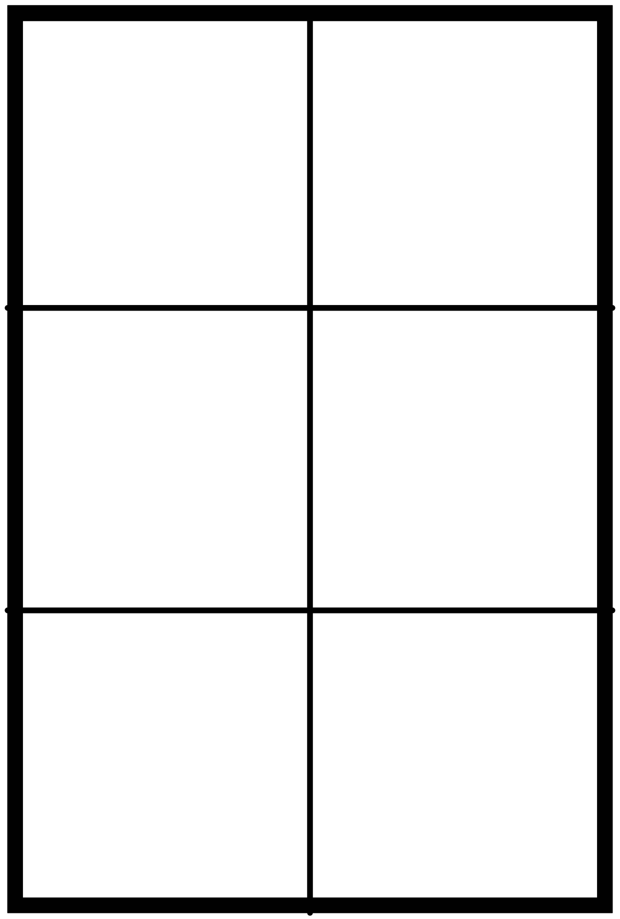
*Much work has been done to derive multiple test procedures that achieve this goal!*

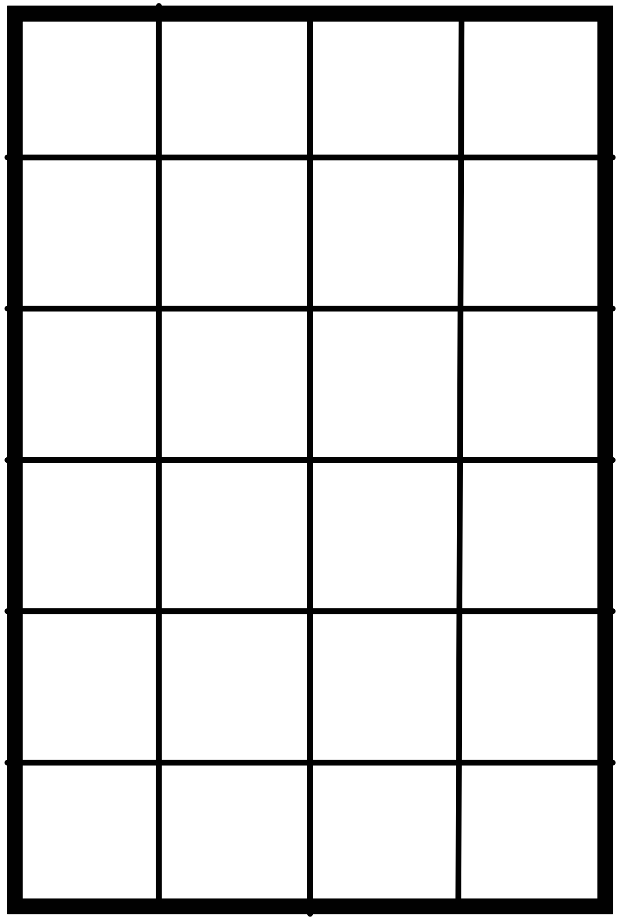
*We address issue from a different as usual point of view ...*

# Our Framework

Consider situation where ...

- multiple hypotheses are to be tested
- there is control at the design stage concerning how many hypotheses will be tested
- overall number of observations is limited by some constant $m$
- there is control at the design stage concerning the allocation of the observations among the hypotheses to be tested





# Some Applications

- Clinical trials with subgroups defined by age, treatment etc.
- Crop variety selection
- Microarrays
- Discrete event systems

# Our Goal

Given

- a maximum overall number of observations,
- a certain multiple test procedure

Maximize (in number $k$ of considered hypotheses):
expected number of correct rejections

# Outline

- Framework of optimization problem
- Optimization w.r.t. a reference alternative
  - Optimum number of hypotheses when controlling the family-wise error (Bonferroni, Bonferroni–Holm, Dunnett)
  - Optimum number of hypotheses when controlling the false discovery rate (Benjamini–Hochberg)
- Optimization w.r.t. a composite alternative
- Classification Procedures

# The Optimization Problem

- Total of $m$ observations and $K$ potential hypotheses pairs available.
- Focus on hypotheses of type $H_{0,i} : \theta_i = 0$ vs. $H_{1,i} : \theta_i > 0$, $(1 \leq i \leq K)$.
- If $k$ hypothesis pairs selected at random, $m/k$ observations available for each hypothesis pair (up to round off differences).
- Choose $k$ to maximize expected number of correct rejections $EN_k$.

# General Observations

- If no correction for multiplicity applied, $k$ as large as possible is often optimal.
- With correction for multiplicity, there is usually a unique optimum $k$.

# Bonferroni Tests

Define

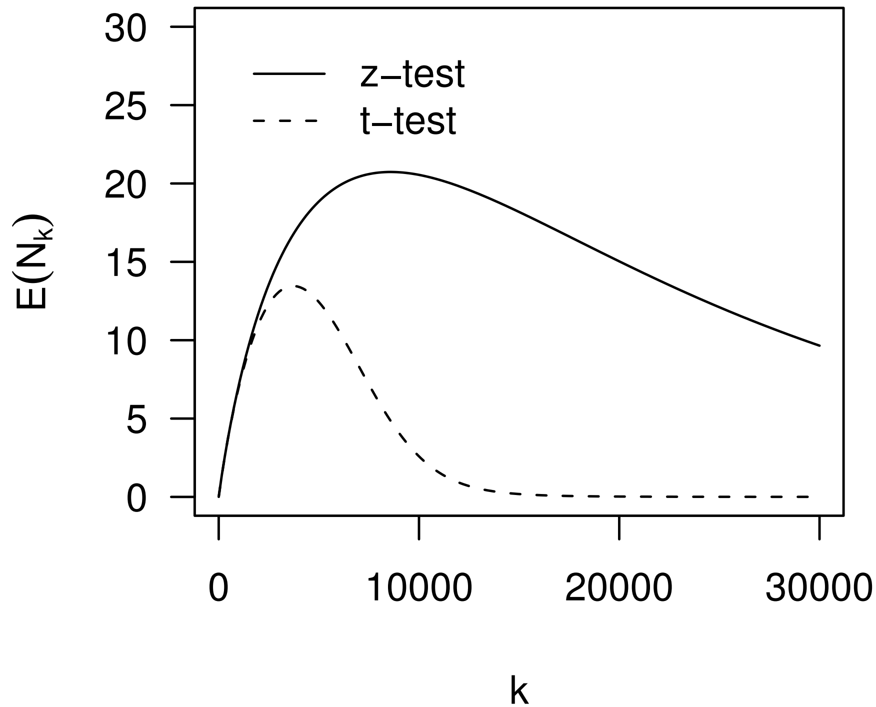
$$\Delta_m := \theta^{(1)} \frac{\sqrt{m}}{\sigma}$$

Then, for normally $N(0, \sigma^2)$ distributed data and one-sided Bonferroni z-tests:

$$E(N_k) = q\,k\left(1 - \Phi_{(\Delta_m/\sqrt{k},1)}(z_{\alpha/k})\right)$$

where $q$ is the expected proportion of incorrect null hypotheses.

# Example: Bonferroni z- and t-tests




*The expected number of correctly rejected null hypotheses for given* $k$ *and the parameters* $m = 100000$, $q = 0.01$, $\alpha = 0.05$, *and* $\theta = \sigma$ *under* $H_1$.

# Optimum Number of Hypotheses

**Theorem:** Define

$$k_m := \frac{\Delta_m^2}{2\log(\Delta_m^2)}.$$

Then, as $m \to \infty$, the optimum number of hypotheses to test is

$$k_m^* = k_m[1 + o(1)],$$

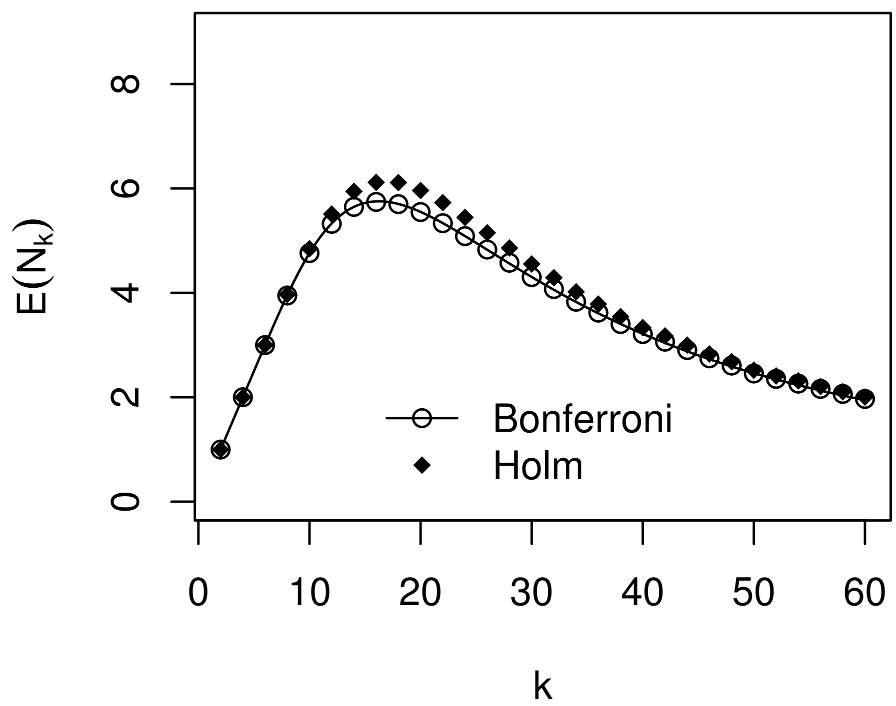
with remainder term being negative.

# Numerical Example

The optimum number of hypotheses $k_m^*$ and the power (in $\%$) to reject an individual incorrect null hypotheses:

| | | $\Delta_m$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 5 | 10 | 20 | 50 | 100 | 1000 |
| | $0.01$ | 3 (57) | 8 (70) | 25 (74) | 124 (76) | 425 (78) | 28908 (82) |
| $\alpha$ | 0.025 | 3 (69) | 9 (71) | 29 (72) | 138 (75) | 469 (77) | 30883 (81) |
| | $0.05$ | 4 (60) | 11 (66) | 33 (70) | 152 (74) | 508 (76) | 32564 (81) |

# Bonferroni–Holm Tests




**Bonferroni vs. Bonferroni–Holm Tests:**

$\theta = 1, m = 200, \alpha = 0.025$, and $q = 0.5$.

# Control of False Discovery Rate

**Benjamini–Hochberg:**

$$FDR = E(\frac{V}{\max(R,1)})$$

Asymptotically equivalent problem (see Genovese and Wasserman (2002)):

$$E(N_k) = q\,k\,\left(1 - \Phi_{(\Delta_m/\sqrt{k},1)}(z_u)\right) \to \max_k,$$

# Benjamini–Hochberg

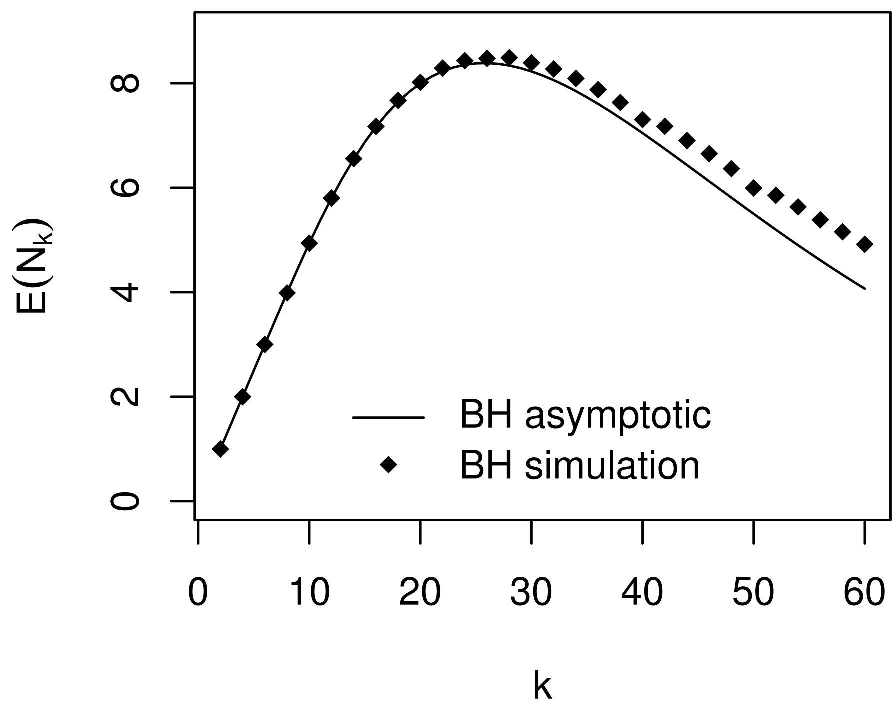
**Theorem:** Asymptotically, the optimum solution is

$$k_m^* = \frac{\Delta_m^2}{\left(z_{u_\beta^*} - z_{\beta u_\beta^*}\right)^2},$$

where $u_\beta^*$ maximizes

$$\frac{u}{\left(z_u - z_{\beta u}\right)^2}.$$

# Asymptotic vs. Simulated Objective Function




The parameters: $\theta = 1$, $m = 200$, $\alpha = 0.025$, and $q = 0.5$.

# t-Tests I

**Bonferroni-tests:**

$$\mathbf{E} N_k^{(t)} = q\,k[1 - F^{(t)}_{m/k,\Delta_m/\sqrt{k}}(t_{\alpha/k,m/k})],$$

with $F^{(t)}_{\nu,\delta}$ non-central t-cdf with $\nu - 1$ df and noncentrality parameter $\delta$, and $t_{\gamma,\nu}$ $1-\gamma$ quantile of standard t-distribution with $\nu - 1$ degrees of freedom.

**Benjamini–Hochberg procedure:**

$$\mathbf{E} N_k^{(t)} = q\,k[1 - F^{(t)}_{m/k,\Delta_m/\sqrt{k}}(t_{u,m/k})].$$

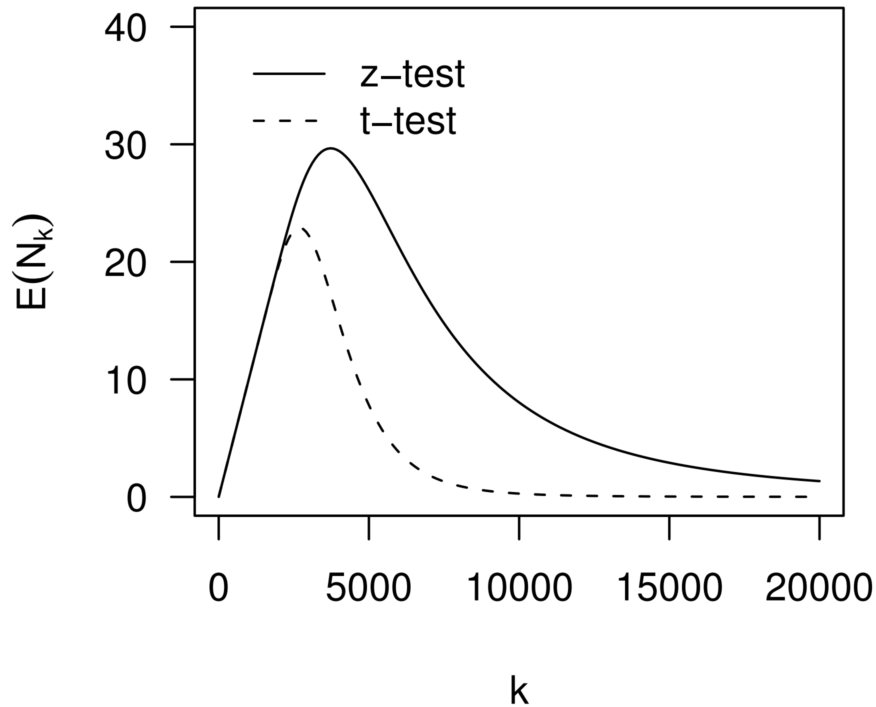
# t-Tests II

**Theorem:** Let $\theta^{(1)} > 0$, and define $\theta_m = \theta/\sqrt{m}$. Assume that

$$\Delta_m = \frac{\theta_m \sqrt{m}}{\sigma} = \frac{\theta^{(1)}}{\sigma}.$$

Then, for $m \to \infty$, the optimum solution for t-tests converges to that for z-tests.

# Possible Rejections for z- and t-Test




Parameters: $m = 100000$, $q = 0.01$, $\alpha = 0.05$ and $\theta^{(1)}/\sigma = 1$.

# Composite Alternatives I

**Bonferroni z-Tests:**

$$EN_k = q\,k \int_0^\infty \left(1 - \Phi(z_{\alpha/k} - \frac{\Delta_m(\theta)}{\sqrt{k}})\right) dF(\theta),$$

where $F$ conditional c.d.f. of $\theta$ given $\theta > 0$,
$q = P(\theta > 0)$, and $\Delta_m(\theta) = \theta \frac{\sqrt{m}}{\sigma}$.

# Composite Alternatives II

**Theorem:** Assume that $F$ is continuous and define

$$k_{m,F} := \frac{m\, d_F^2/\sigma^2}{2\log(m\, d_F^2/\sigma^2)},$$

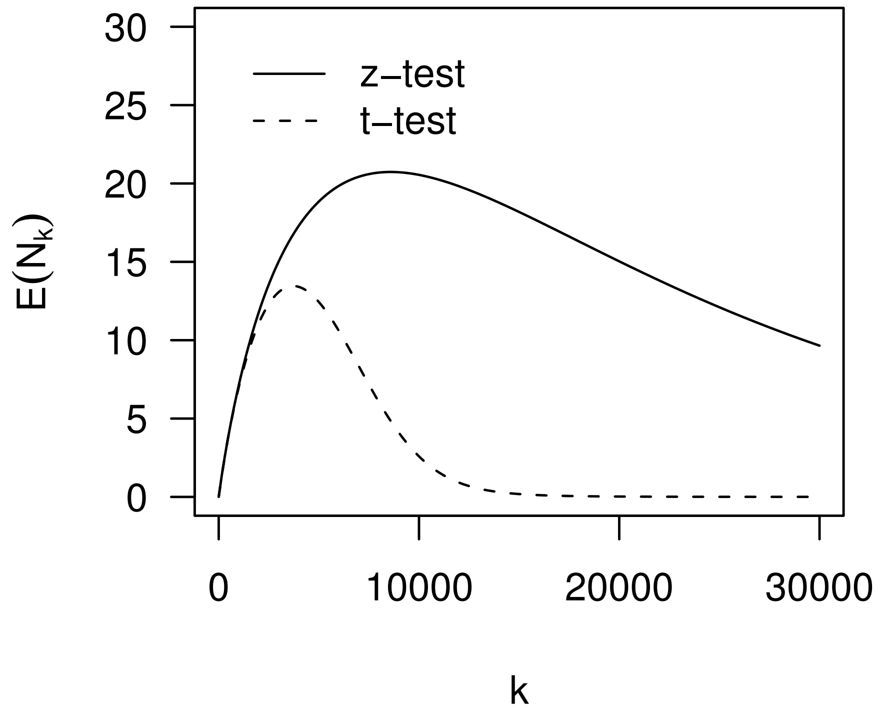
where $d_F$ maximizes

$$d^2[1 - F(d)].$$

Assuming that $d^2(1 - F(d)) \to 0$ as $d \to \infty$, optimum solution $k_{m,F}^*$ satisfies

$$k_{m,F}^* = k_{m,F}(1 + o(1)).$$

# Composite Alternatives III




Parameters: $m = 100000$, $q = 0.01$, $\alpha = 0.05$.
Effect size under alternative $N(0, 1.2)$ distributed.

# Composite Alternatives IV

Similar result can be obtained for Benjamini–Hochberg procedure ...

# Classification Procedures I

Classification between $\theta = \theta_0$ and $\theta = \theta_1$
Minimize

$$k\left(w_1\, q\, [1 - g_k(\theta_1)] + w_0\,(1 - q) g_k(\theta_0)\right),$$

with $g_k(\theta)$ probability of deciding for $\theta_{(1)}$ under $\theta$.
For fixed $k$, problem equivalent to maximizing

$$U(k) = k\left(w_1\, q\, g_k(\theta_1) - w_0\,(1 - q) g_k(\theta_0)\right).$$

# Classification Procedures II

**Theorem:** For Bayes-classifier, normal data and $r = w_0\,(1-q)/(w_1\,q)$:
If $r > 1$, then optimum $k$ satisfies

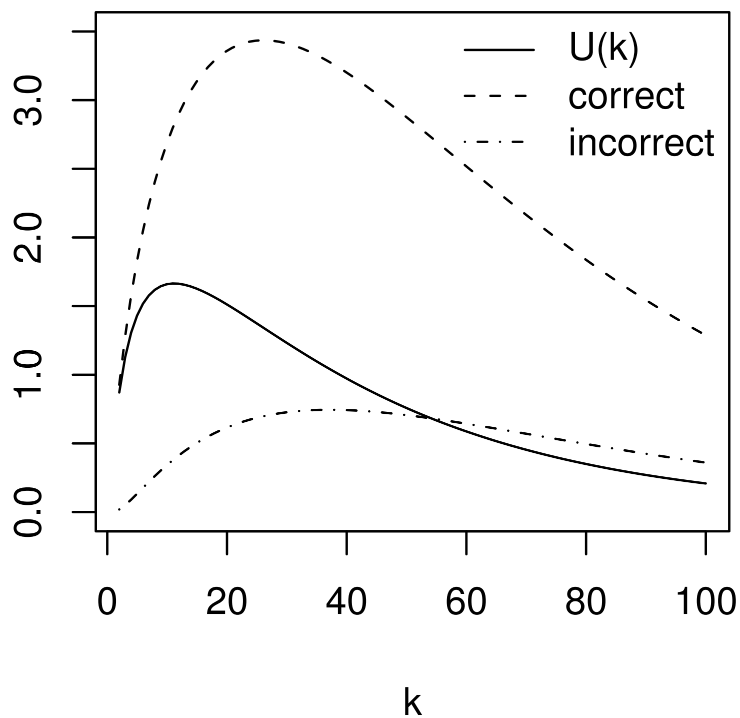
$$k = \left(\frac{\Delta_m}{x_r}\right)^2,$$

where $x_r$ is the solution of

$$0 = x\,\varphi[x - c(r,x)]/2 - \Phi[x - c(r,x)] + r\,\Phi[-c(r,x)],$$

with $c(r,x) = \log(r)/x + x/2$, and
$\Delta_m = \theta_1 - \theta_0\sqrt{m}/\sigma.$

# Objective Function U(k)




Parameters $m = 100$, $q = 0.5$, $w_0 = 3$, $w_1 = 1$, and $\theta^{(1)}/\sigma = 1/2$.

# Summary

- Given a limited maximum number of observations, the number of possible rejections depends considerably on the number of considered hypotheses.
- With a good design involving an appropriate allocation of the observations to the hypotheses, a lot more can be gained than by using a more sophisticated multiple test procedure.
- For more details see: *On the Optimum Number of Hypotheses to Test when the Number of Observations is Limited. (Futschik & Posch (2005))*