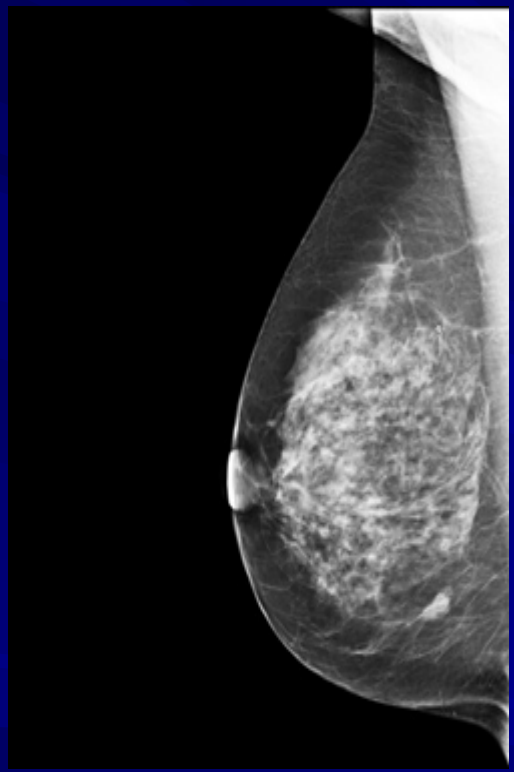
Exact calculations of expected power for the Benjamini-Hochberg procedure

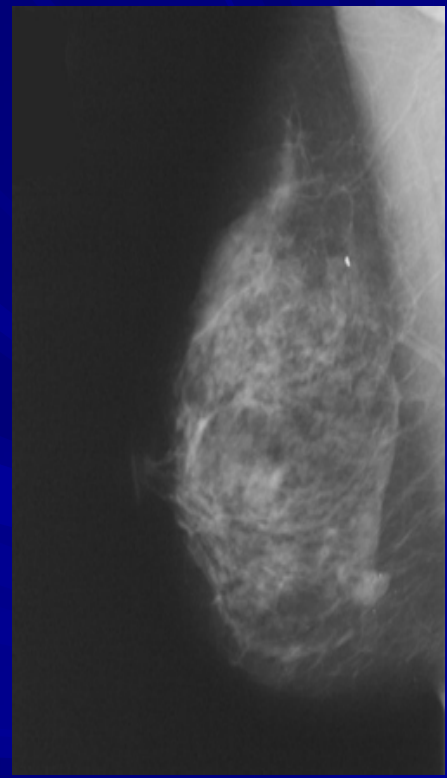
DH Glueck[1], A Karimpour-Fard[1], L Hunter[1], J Mandel[1] and KE Muller[2]

[1]University of Colorado at Denver and Health Sciences Center
and
[2]University of Florida

Which is better, digital or film mammography?



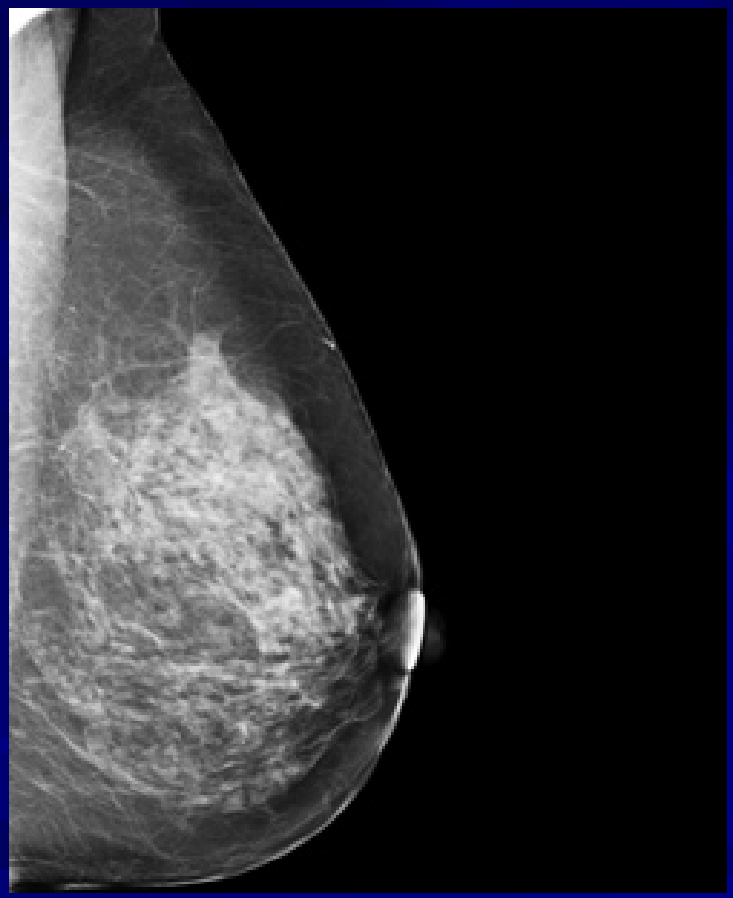
Full Field Digital Mammogram



Screen-Film Mammogram

http://www.hologic.com/wh/cc-050701.htm

Mammography



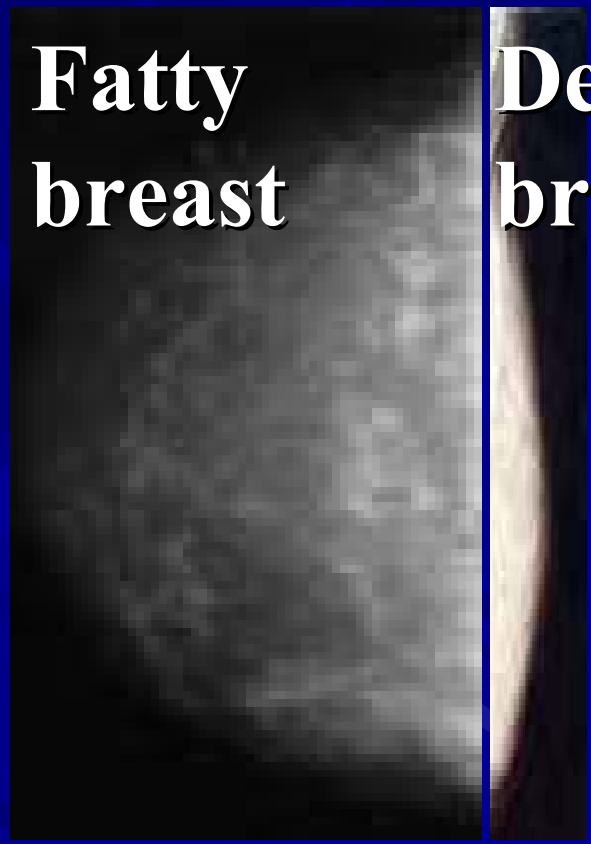
http://www.hologic.com/wh/whccl.cfm

- A screening technique for early detection of breast cancer
- An x-ray of the breast
- **Screen film mammography** uses film
- **Digital mammography** uses a digital detector

Subgroup analysis

Four groups

- **Young women with dense breasts**
- **Young women with fatty breasts**
- **Old women with dense breasts**
- **Old women with fatty breasts**



http://visualsonline.cancer.gov/details.cfm?imageid=2696

Problem

- Many groups
- Many hypotheses
- Multiple testing problem

Study design

- Control false discovery rate using Benjamini and Hochberg procedure
- Use exact methods to calculate power

Multiple comparisons

- Benjamini & Hochberg (1995)
- Benjamini & Yekutieli (2001)
- Storey (2002) JRSS
- Sandrine Dudoit (2003)

Power for multiple comparisons

- Efron, Storey and Tibshirani (2001) JASA
- Lee and Whitmore (2002)
- Finner & Roters (2002)
- Sarkar (2002, 2004, 2006)
- Zien, Fluck, Zimmer and Lengauer (2002)
- Lehmann &Romano (2005), Romano &Shaikh (2006)

Benjamini and Hochberg procedure

1. *independent* tests and $\alpha_* \in (0,1)$ is control target.

2. m statistics $\{t_i\}$, p-values, $\{p_i\}$.

3. Rank: $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(m)}$

4. Find largest k such that $p_{(k)} \leq k \cdot \alpha_* / m$.

5. Reject hypotheses with k smallest p-values

$\Rightarrow$ False Discovery Rate $\leq \alpha_*$

Decisions for an experiment

True State	Decision: Reject	Decision: Accept	
H_0 True	j	$n - j$	n
H_0 False	$k - j$	$(m - n) - (k - j)$	$m - n$
	k	$m - k$	m

Expected power

$$\mathcal{E}\left[\frac{(K-J)}{(m-n)}\right] =$$

$$\sum_{k=0}^{m}\sum_{j}\left[\frac{(k-j)}{(m-n)}\right]\Pr\{(K=k)\cap(J=j)\}$$

Exact expected power

Fix

- # of hypotheses
- # of null and alternative hypotheses
- distributions of the test statistics under the null and the alternative

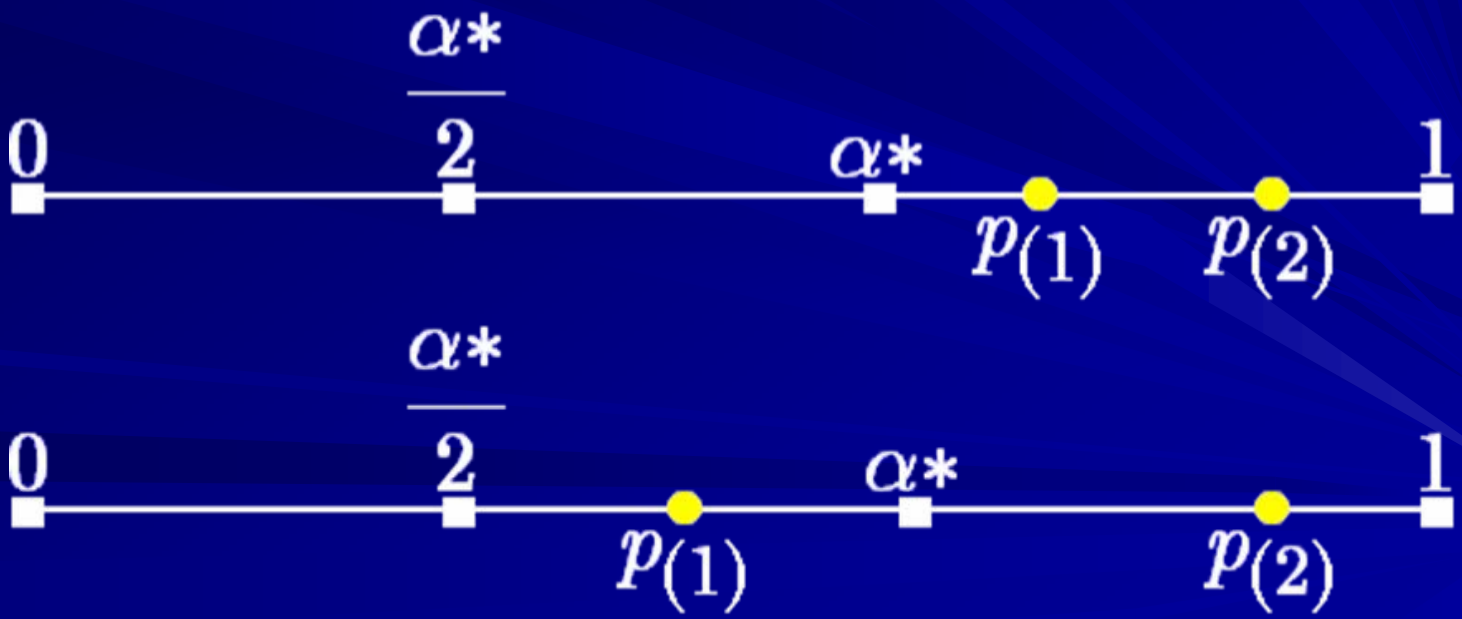
Exact expected power

- Partition the rejection regions into disjoint subsets
- Calculate the probability of each subset
- Find the joint probability distribution of the number of total and false rejections
- Use this joint distribution to give explicit formulas for the expected power

Partitioning rejection regions

$$m = 2$$

$$k = 0$$



Number in partition is Catalan

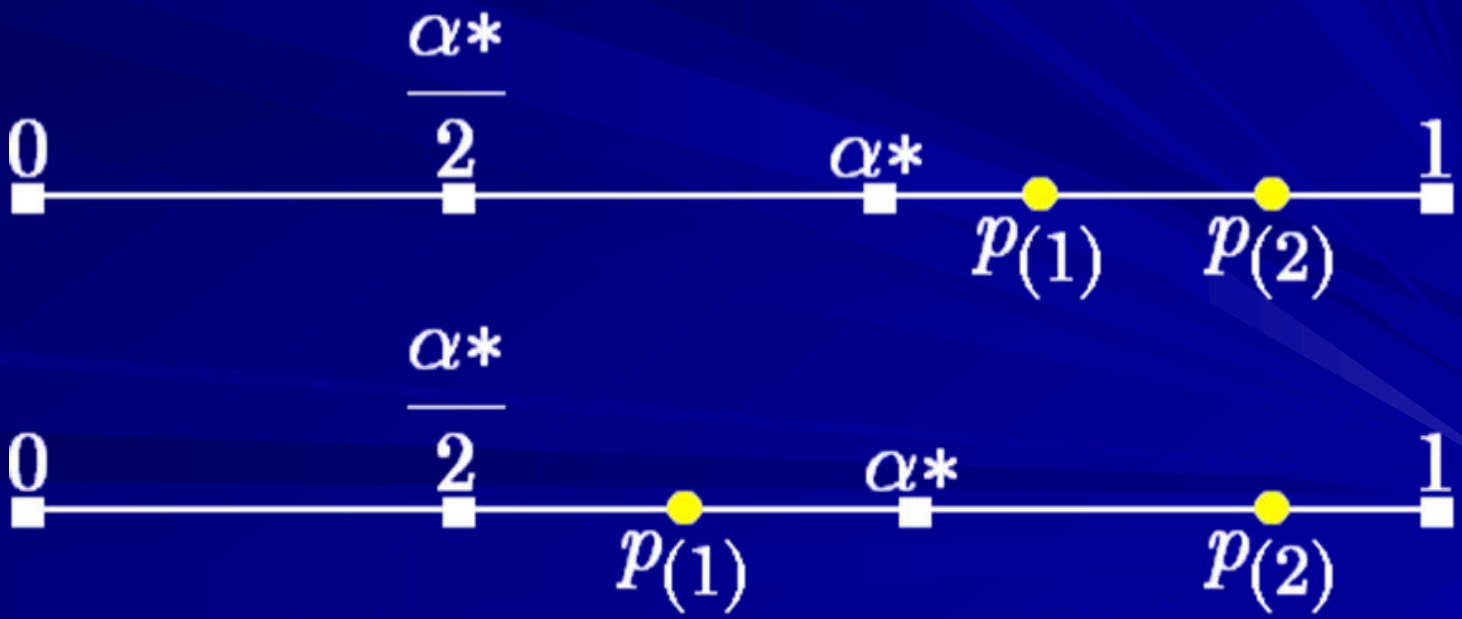
$$c_k = \frac{[2 \cdot (m-k)]!}{[(m-k)!(m-k)!(m-k+1)]}$$

Glueck et al. 2006a

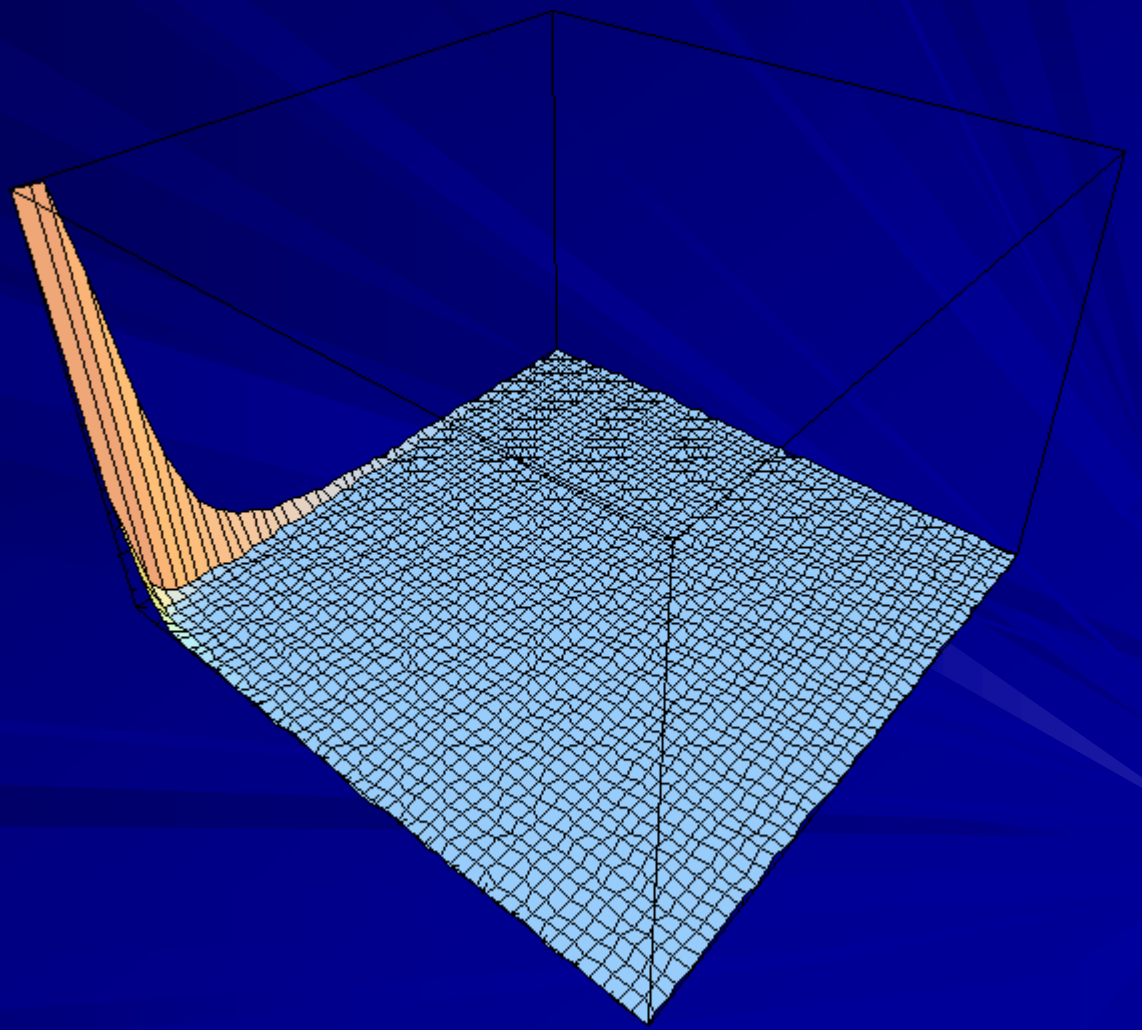
Partitioning rejection regions

$$m = 2$$

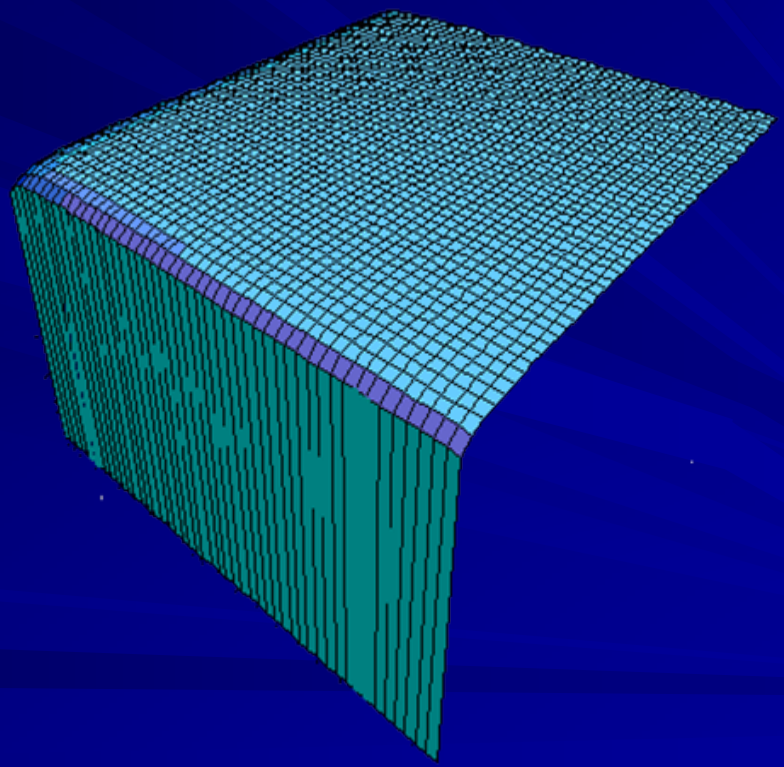
$$k = 0$$



Joint density of p-values for two one sample z tests



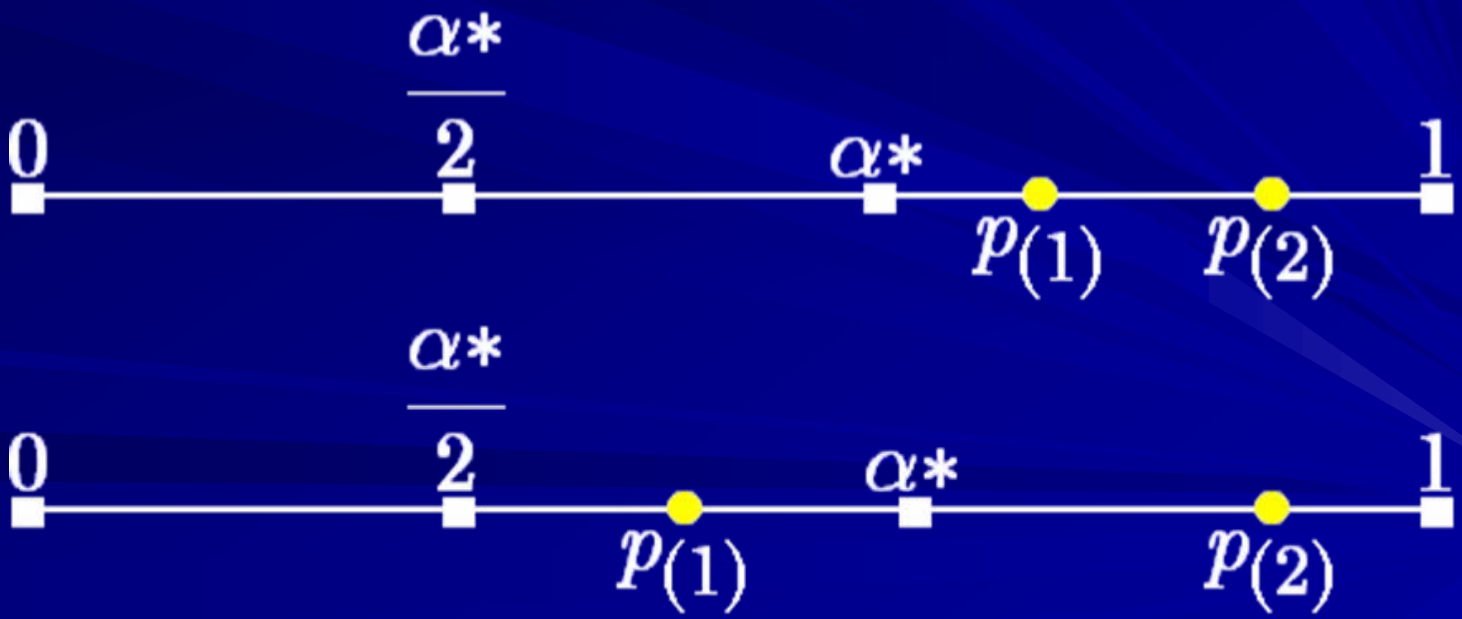
Joint cumulative distribution function of p-values for two one sample *z* tests.



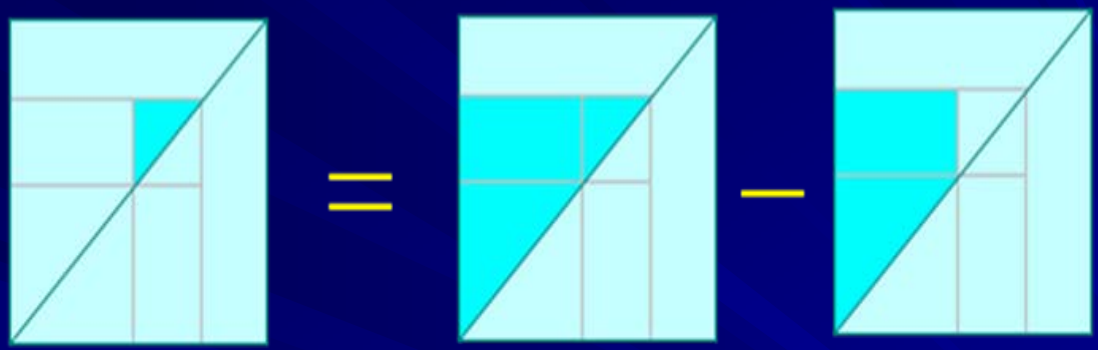
Partitioning rejection regions

$$m = 2$$

$$k = 0$$



Area in ordered space



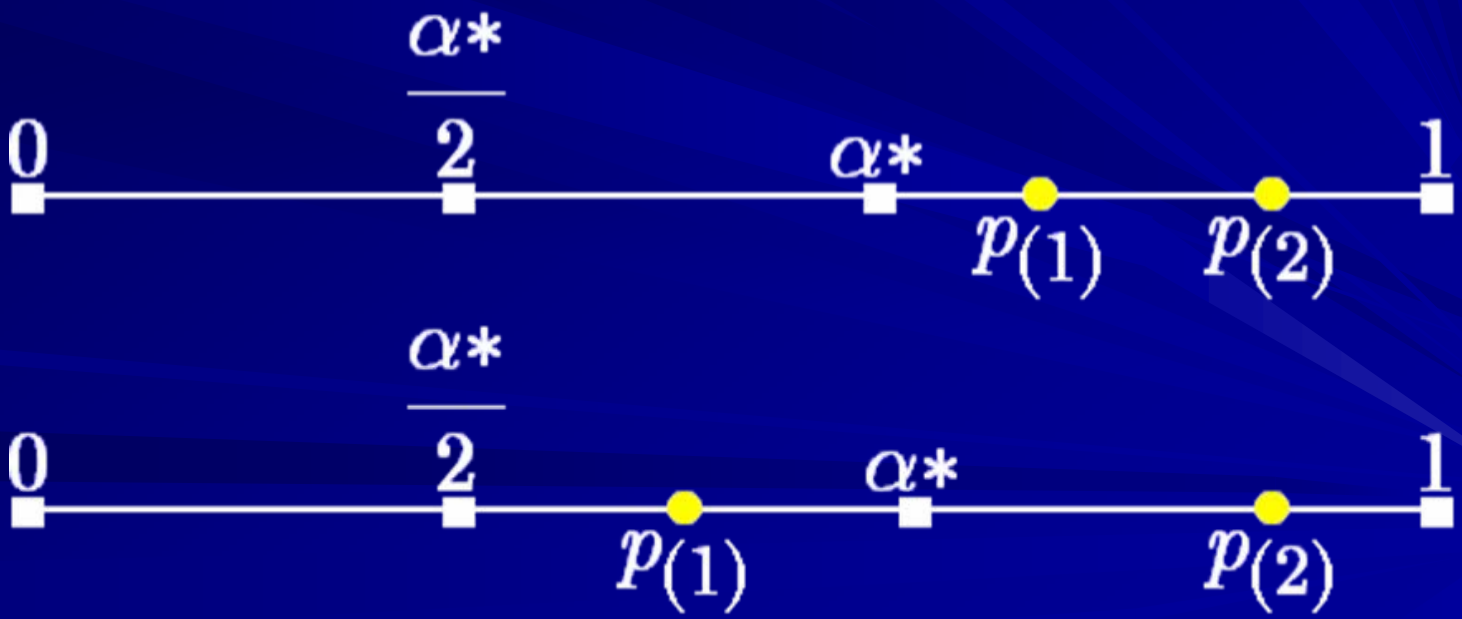
Glueck 2006b

Glueck, 2007a

Partitioning rejection regions

$$m = 2$$

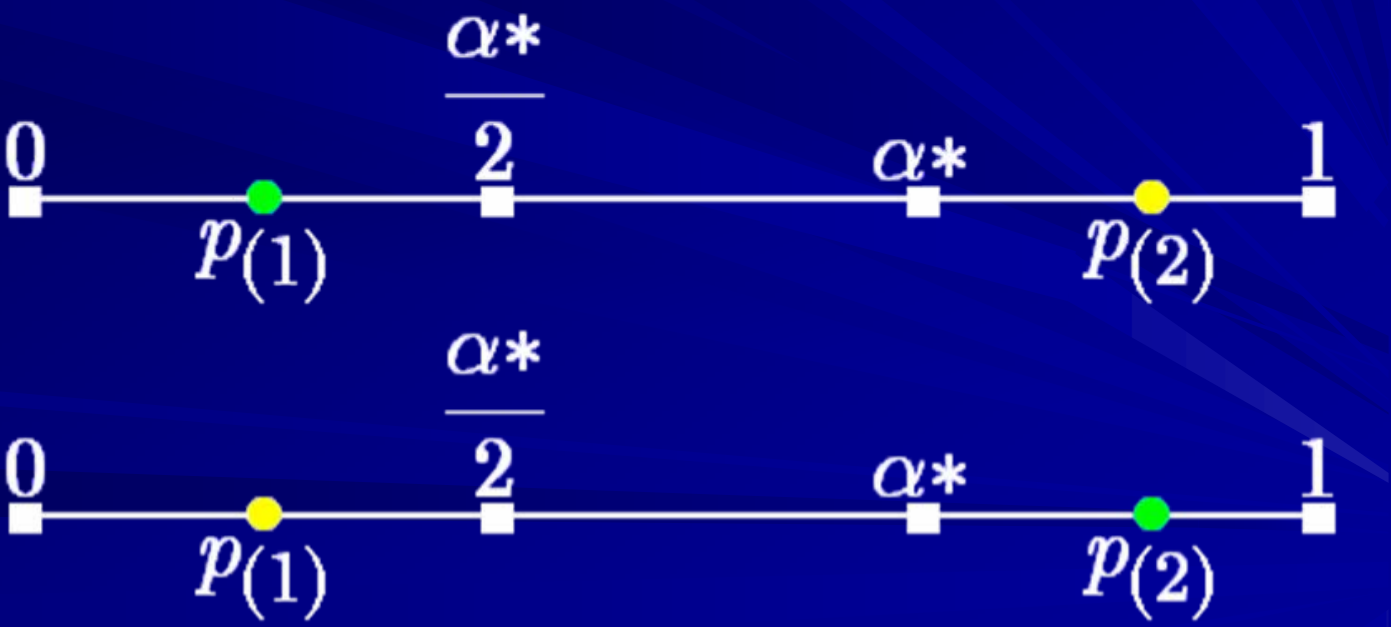
$$k = 0$$



Joint probability of *K* and *J*

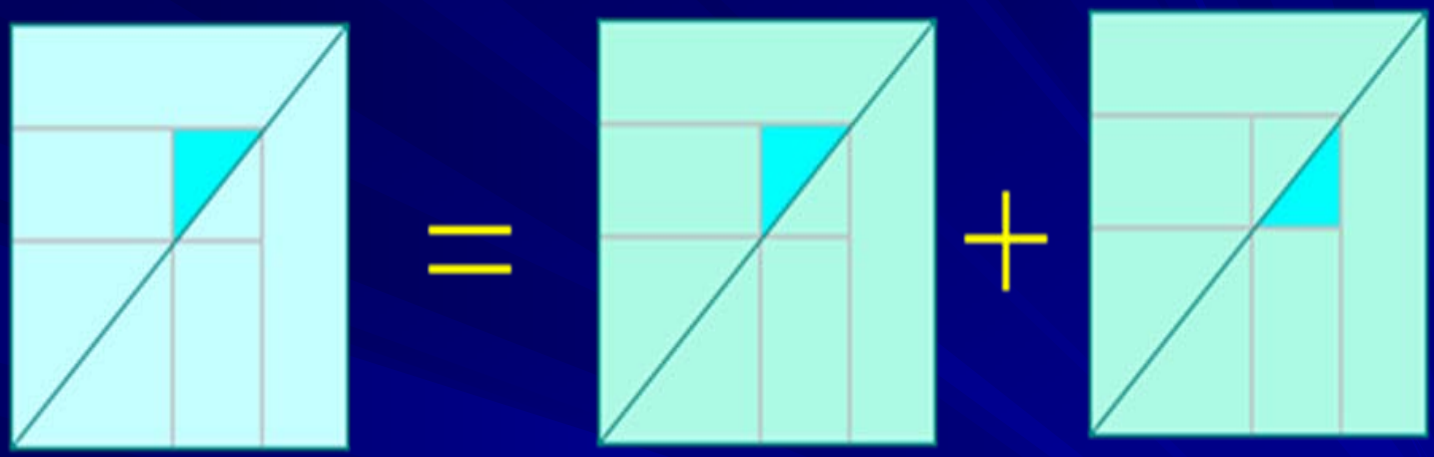
$$m = 2$$

$$k = 1$$



Glueck, 2007b

Probabilities of rejecting nulls



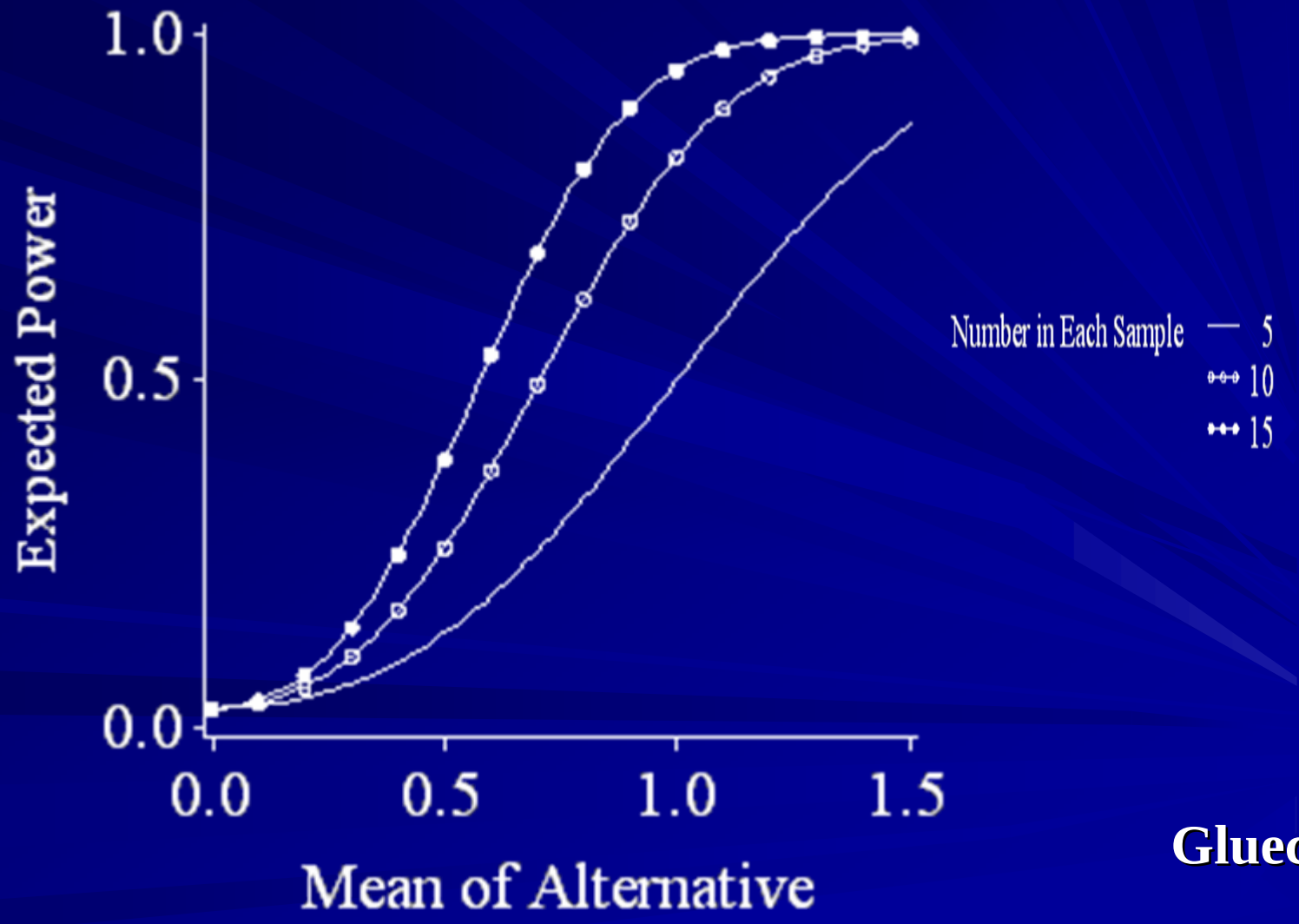
Glueck, 2007b

Expected power

$$\mathcal{E}\left[\frac{(K-J)}{(m-n)}\right] =$$

$$\sum_{k=0}^{m}\sum_{j}\left[\frac{(k-j)}{(m-n)}\right]\Pr\{(K=k)\cap(J=j)\}$$

Exact power



Glueck 2007c

References

- Glueck, D H., Muller, Keith E., Karimpour-Fard, Anis, Hunter, Lawrence. (2006a). Expected Power for the False Discover Rate with Independence. (in review) .
- Glueck, D. H., Karimpour-Fard, A., Mandel, J. and Muller, K.E. (2006b). On the probability that order statistics fall in intervals. (in review) .
- Glueck, D. H., Karimpour-Fard, A., Mandel, J. , Hunter, L. and Muller, K.E. (2007a) ,Fast computation by block permanents of cumulative distribution functions of order statistics from several populations. arXiv:0705.3851 (in review) .
- Glueck, D. H., Karimpour-Fard, A., Mandel, J. and Muller, K.E. (2007b). On probabilities for separating sets of order statistics. (in review) .
- Glueck, D. H., Karimpour-Fard, A., Hunter, L. , Mandel, J. , and Muller, K.E. (2007c) , Calculating power for the Benjamini-Hochberg procedure. (in progress).

Acknowledgement

- Glueck was supported by NCI K07CA88811.
- Hunter was supported by NIAAA 1U01 AA13524-02 and NCI 5 P30 CA46934-15
- Mandel was supported by NSF-CMS0325314.
- Muller was supported by NCI P01 CA47 982-04, NCI R01 CA095749-01A1 and NIAID 9P30 AI 50410.