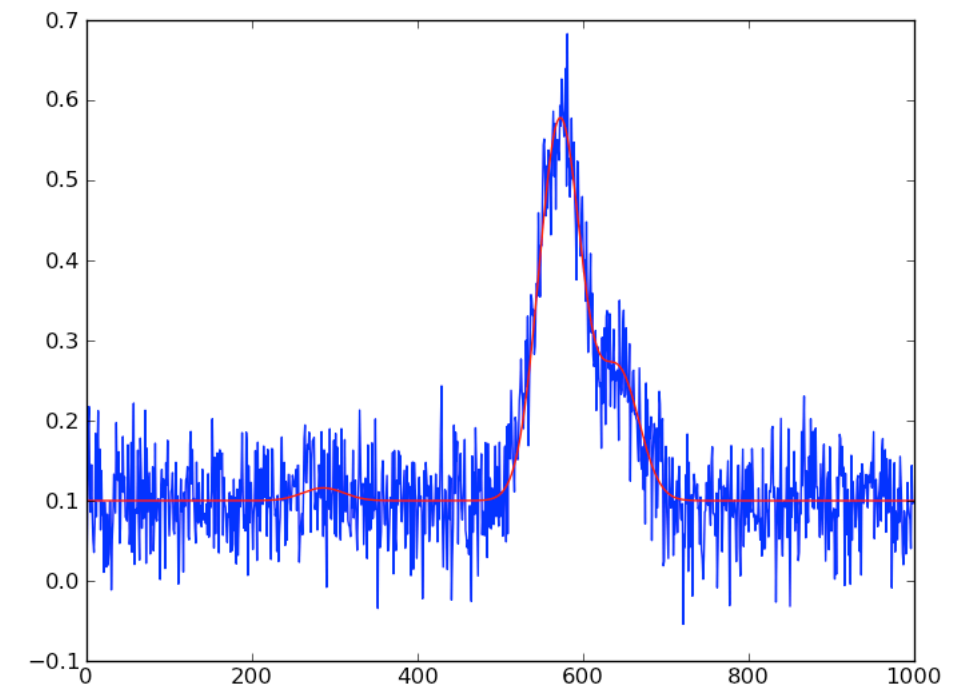
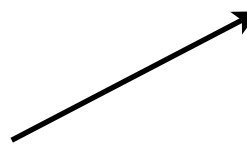
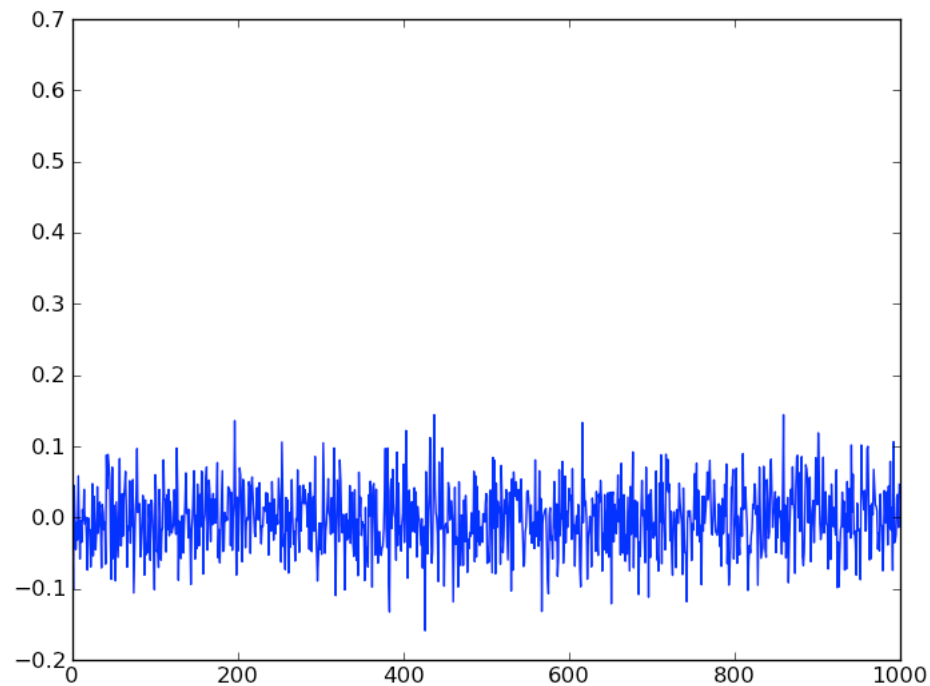
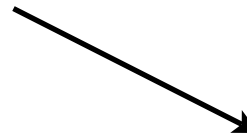
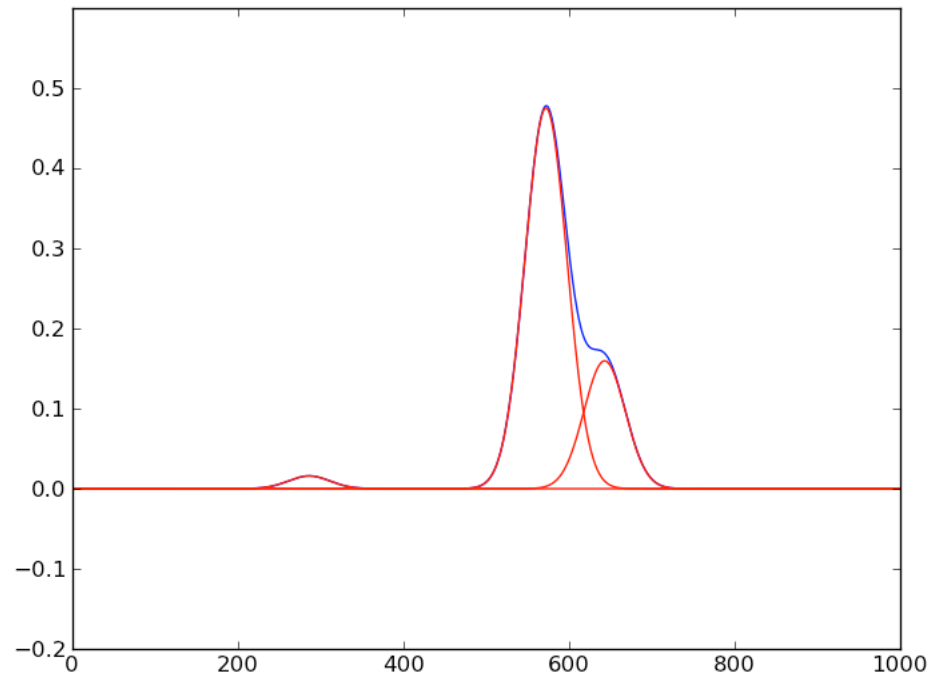


AS4501-CC4501

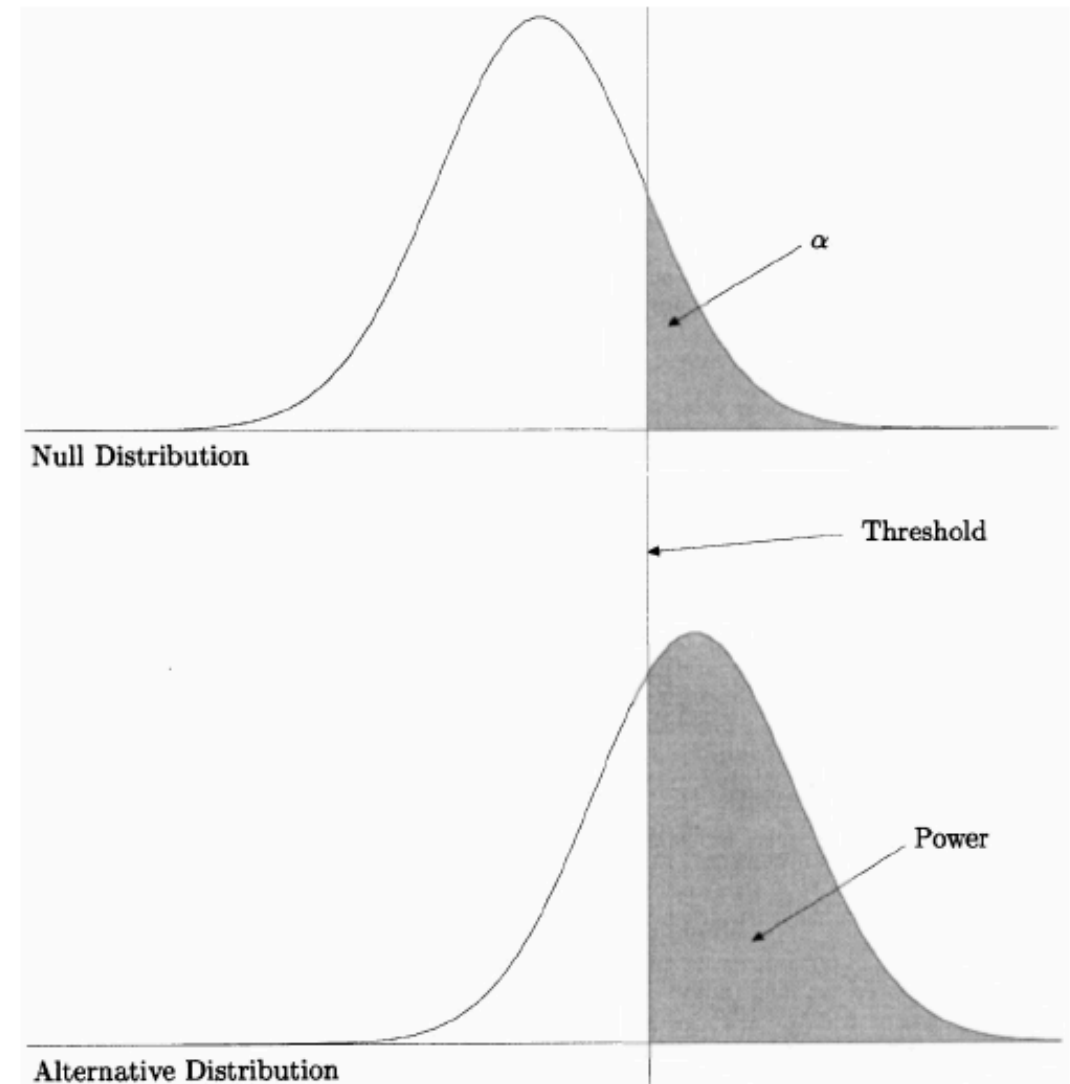
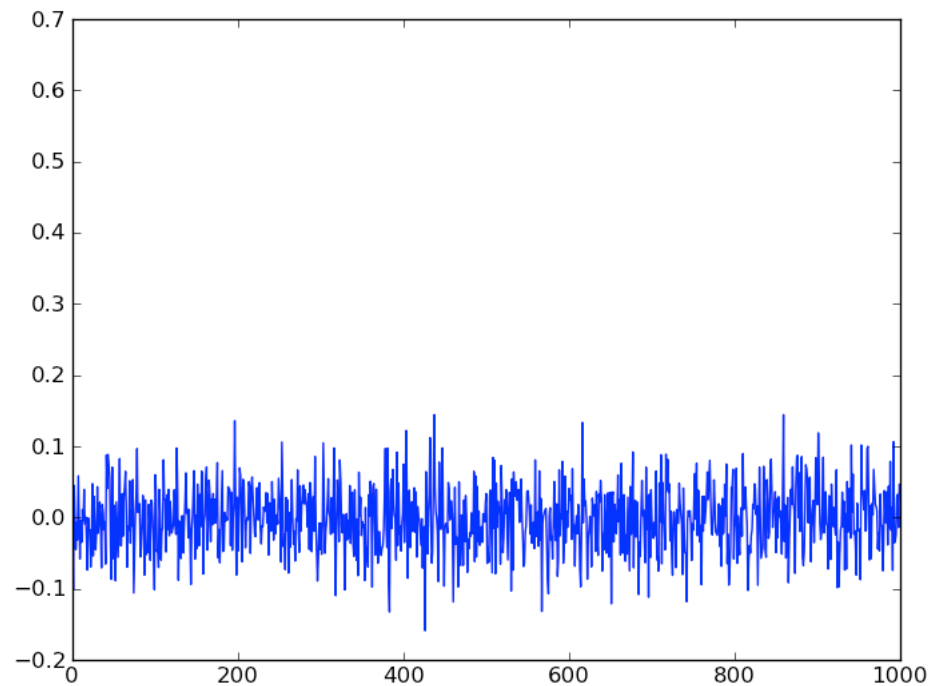
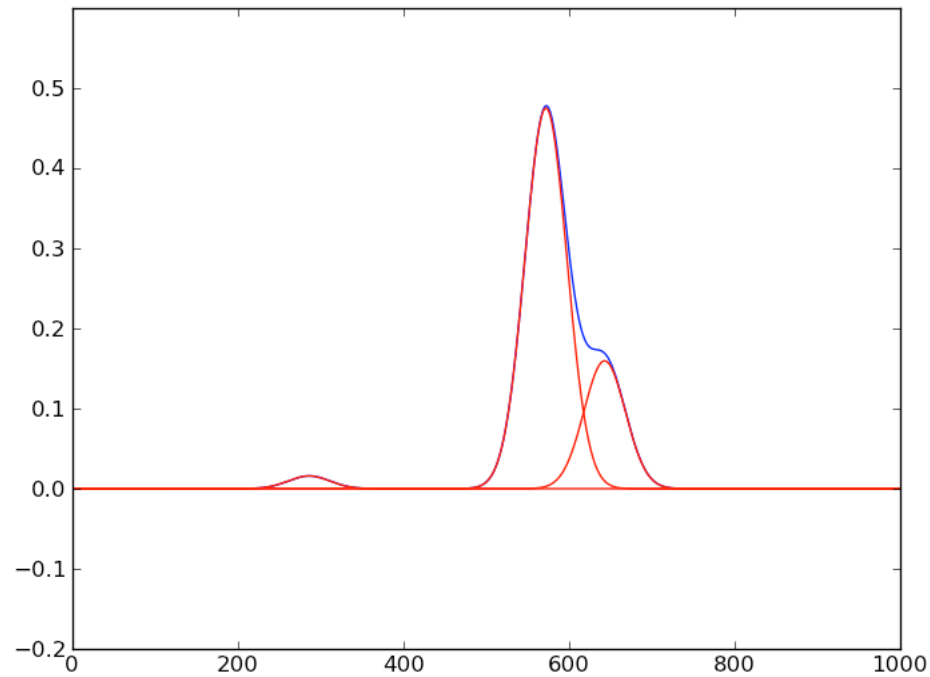
Thresholding vs False Discovery Rate

Guillermo Cabrera (AURA-CTIO / CMM & DCC, University of Chile)

Señal y Ruido



Señal y Ruido



Detección

Miller et.al. 2001

- **Null Hypothesis:** pixel es background
- **Alternative Hypothesis:** pixel pertenece a una fuente
- Errores:
 - Detectar incorrectamente un pixel del background como fuente (false discovery)
 - Detectar incorrectamente un pixel de una fuente como background

Detección

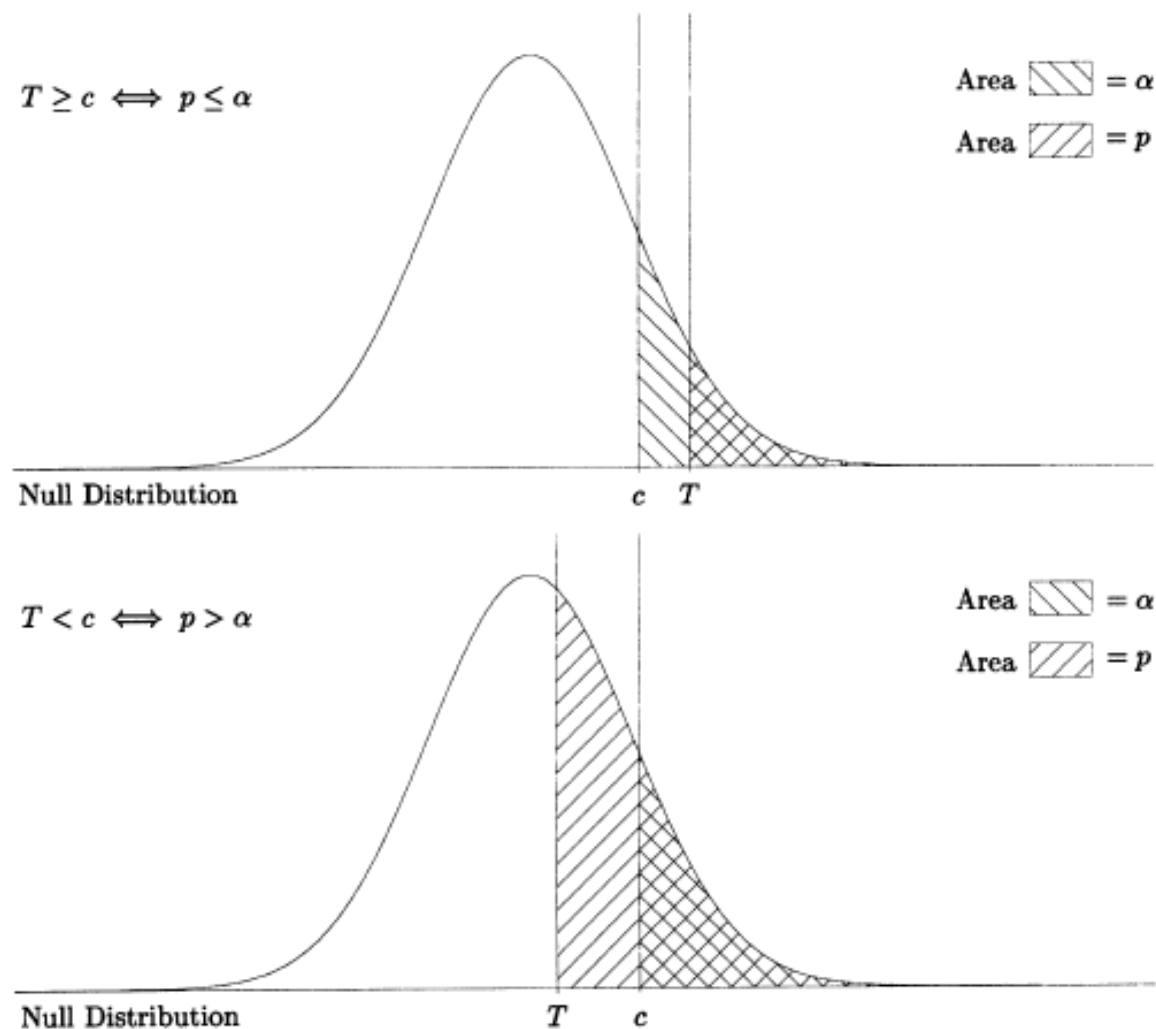
Miller et.al. 2001

TABLE 1

SUMMARY OF OUTCOMES IN MULTIPLE TESTING

Condition	Reject Null	Maintain Null	Total
Null true	$N_{\text{reject null true}}$	$N_{\text{maintain null true}}$	$N_{\text{null true}}$
Null false	$N_{\text{reject null false}}$	$N_{\text{maintain null false}}$	$N_{\text{null false}}$
Total	N_{reject}	N_{maintain}	N

Deteccion



T: Test statistics
c: critical threshold

$\alpha = 0.05$ is equivalent to using a “2 σ ”

- k sigma:

- Asegura $\langle \text{PFD} \rangle = \alpha$
- PFD: Proportion of False Discoveries $N_{\text{null true}}^{\text{reject}}/N$

- Bonferoni:

- Valor de significancia dependiente de N $\alpha'_N = \alpha/N$
- Asegura $\langle \text{AFD} \rangle \leq \alpha$
- AFD: Any False Discoveries?
- 1 si $N_{\text{null true}}^{\text{reject}} > 0$
- 0 si $N_{\text{null true}}^{\text{reject}} = 0$

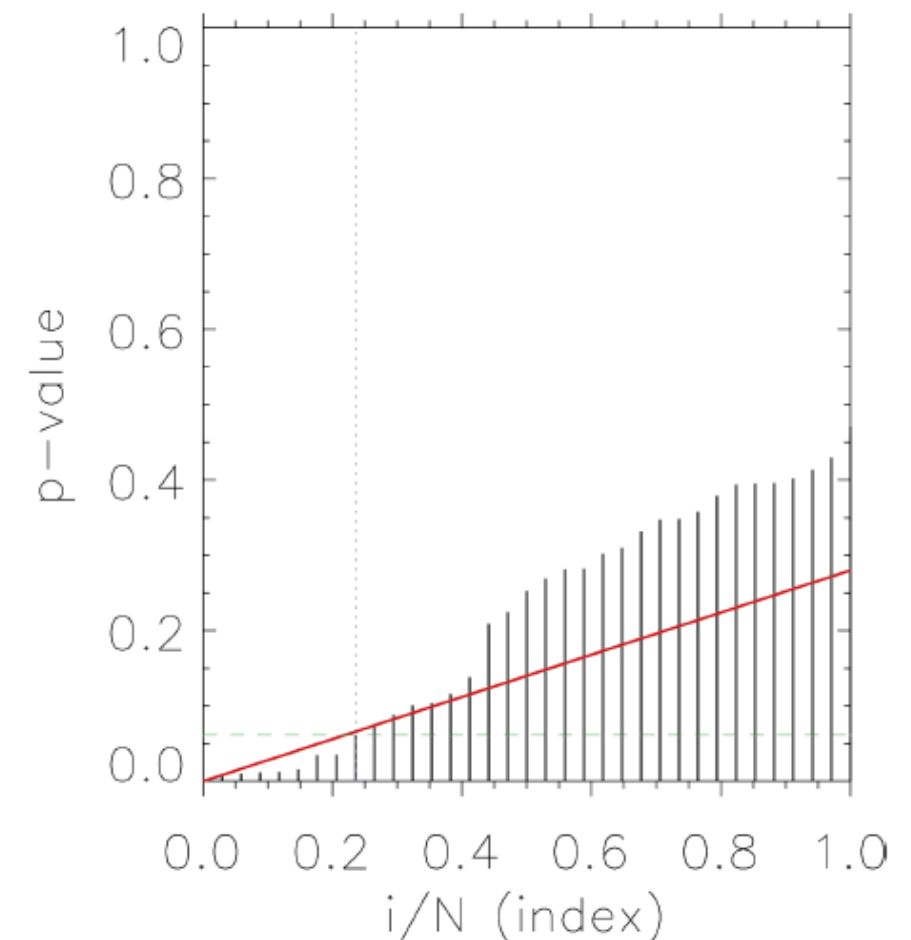
$\langle \rangle$: Esperanza

False Discovery Rate

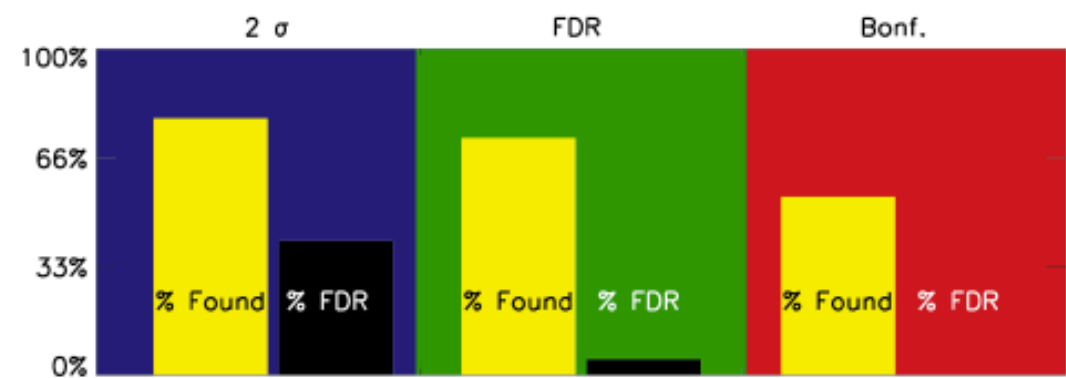
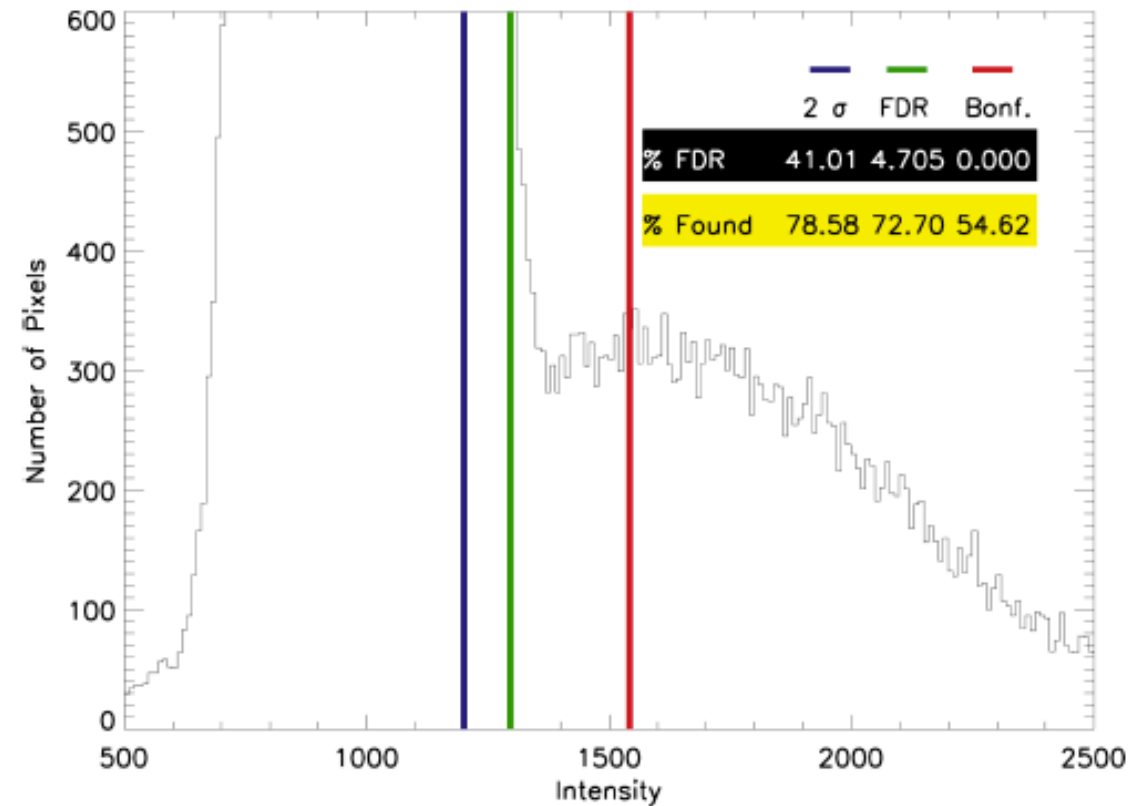
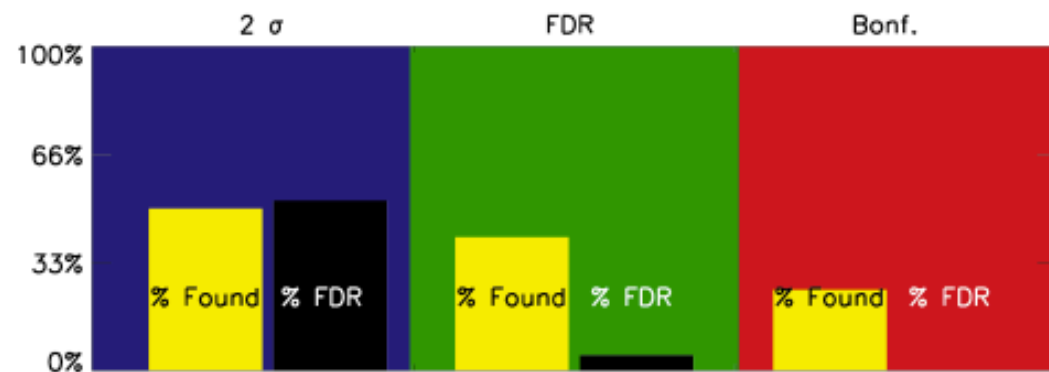
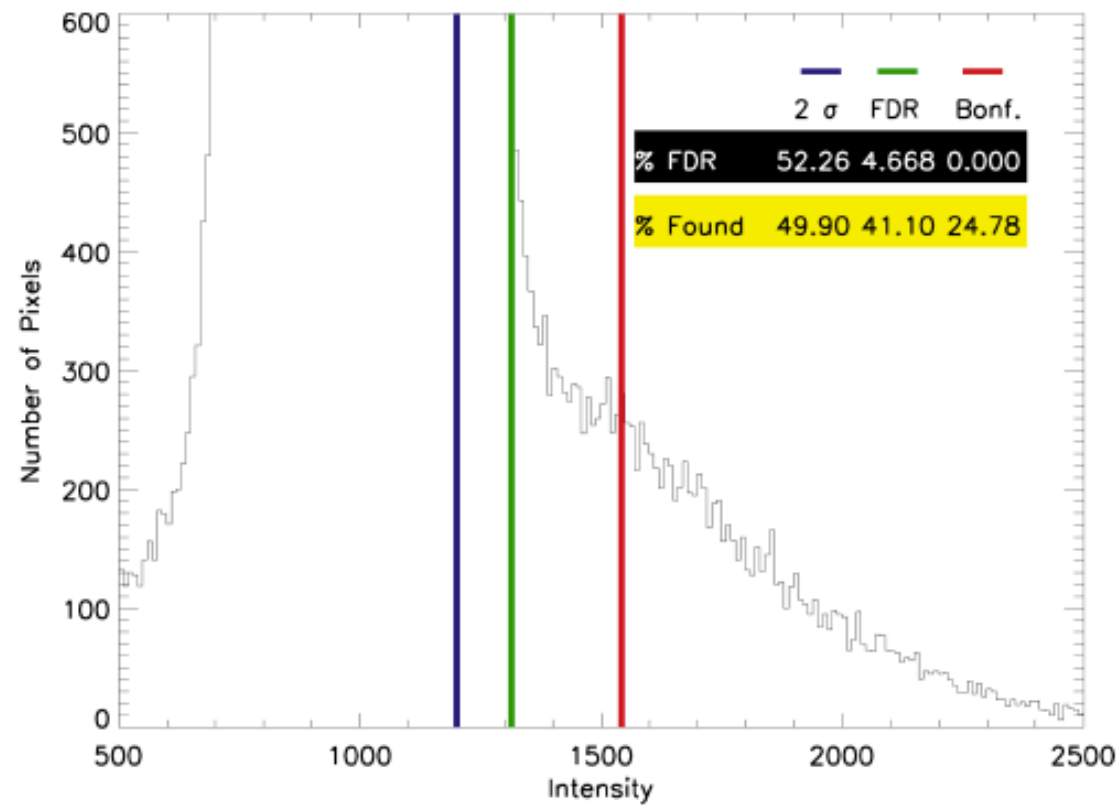
$$\text{FDR} = \frac{N_{\text{null true}}^{\text{reject}}}{N^{\text{reject}}} = \frac{N_{\text{null true}}^{\text{reject}}}{N_{\text{null true}}^{\text{reject}} + N_{\text{null false}}^{\text{reject}}}$$

Asegura $\langle \text{FDR} \rangle \leq \alpha$

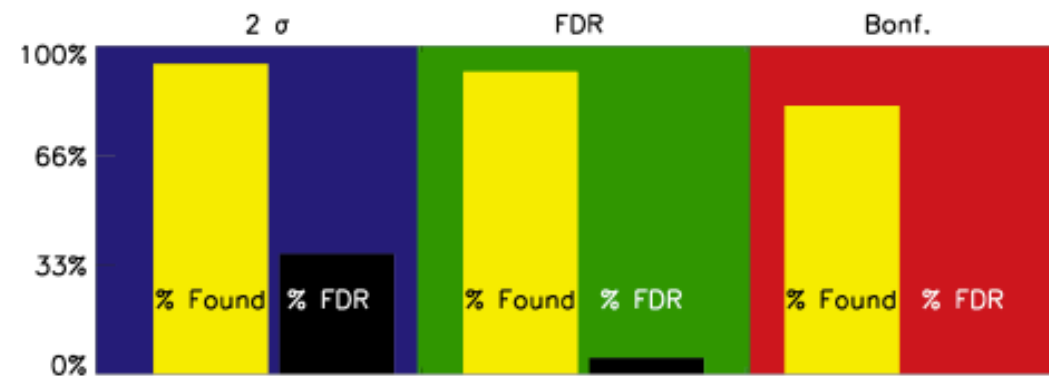
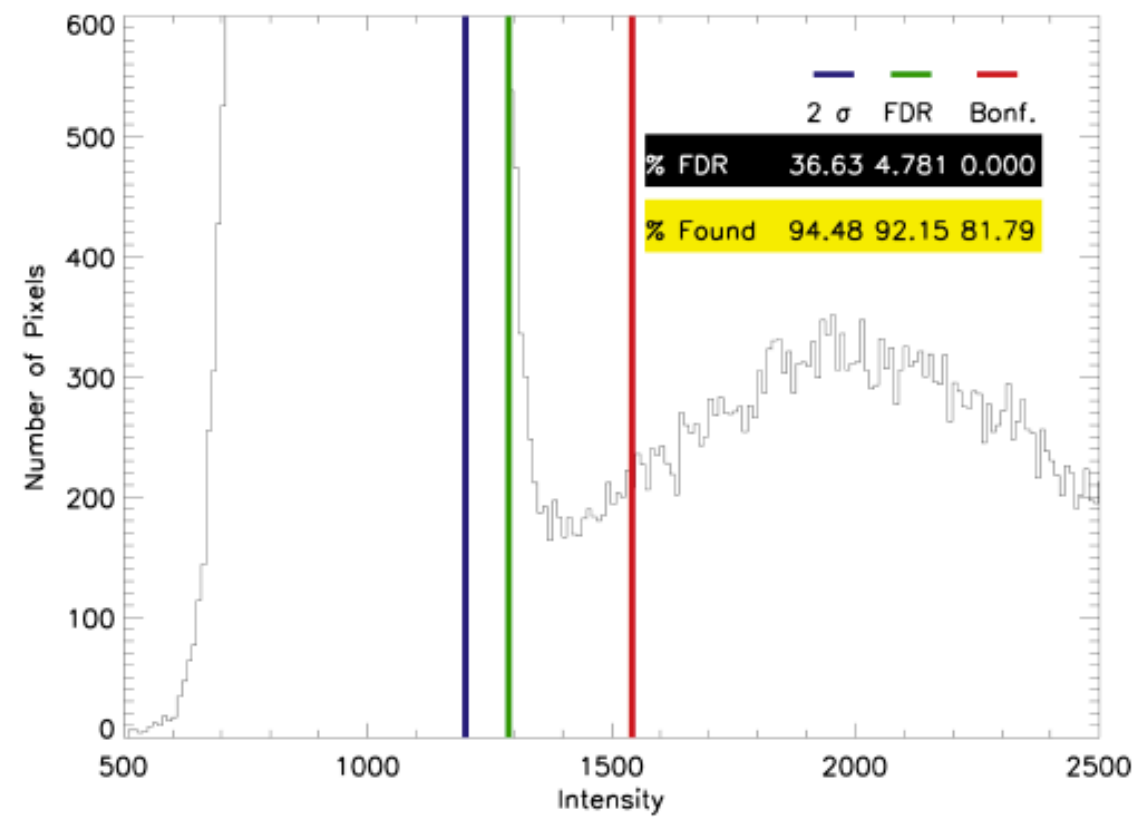
- Seleccionar $0 \leq \alpha \leq 1$
- Sean $P_1, \dots, P_N$ los p-values de las N pruebas (píxeles, por ejemplo).
- Sea $d = \max \left\{ j : P_j < \frac{j\alpha}{c_N N} \right\}$, $c_N = 1$ para muestra independientes, $c_N = \sum_{i=1}^N i^{-1}$ sino.
- Rechazar todas las muestras con p-value menor que P_d



Ejemplos



Ejemplos



Referencia

CONTROLLING THE FALSE-DISCOVERY RATE IN ASTROPHYSICAL DATA ANALYSIS

CHRISTOPHER J. MILLER,¹ CHRISTOPHER GENOVESE,² ROBERT C. NICHOL,¹ LARRY WASSERMAN,²
ANDREW CONNOLLY,³ DANIEL REICHART,⁴ ANDREW HOPKINS,³
JEFF SCHNEIDER,⁵ AND ANDREW MOORE⁵

THE ASTRONOMICAL JOURNAL, 122:3492–3505, 2001 December