Confidence Sets Controlling for False Discovery Rate

Howon Ryu

February 15, 2023

1 Multiple Testing

In a testing setting where m null hypotheses $\{H_1, \ldots, H_m\}$ are tested simultaneously, the following contingency table can be constructed according to the claimed significance and the true state (H_0 true or false) of the hypothesis:

	Not Significant	Significant (rejected)	
H_0	TN	FP	m_0
H_1	FN	TP	$m - m_0$
	$m - R$	R	m

In the above, m represents the total number of hypotheses tested, R represents the number of rejected (claimed significant) hypotheses, and m_0 represents the number of hypotheses that are truly not significant (true H_0). R and m are observable. Let $p_1, \ldots, p_m$ be the p-values associated with the hypotheses $H_1, \ldots, H_m$, and let I_0 be the set of all true H_0 hypotheses. Then, for $i \in I_0$, $\mathbb{P}(p_i < \alpha) < \alpha$, $0 \leq \alpha \leq 1$.

1.1 Family-Wise Error Rate

Family-wise error rate (FWER) is defined as:

$$\mathbb{P}(FP \geq 1)$$

which signifies the probability of making one or more Type I error.

- Bonferroni correction: Controlling $\mathbb{P}(FP \geq 1) < \alpha$ by rejecting H_i if $p_i < \frac{\alpha}{m}$, as $\mathbb{P}(FP \geq 1) \leq \mathbb{E}(FP) \leq \frac{m_0 \alpha}{m} \leq \alpha$

1.2 False Discovery Rate

False discovery rate (FDR) is defined as:

$$\mathbb{E}\left(\frac{FP}{R} | R > 0\right) \mathbb{P}(R > 0) = \mathbb{E}\left(\frac{FP}{\max(1, R)}\right)$$

which aims to control the expected proportion of false positives out of all the significantly claimed tests.

- Benjamini-Hochberg(BH) procedure: The BH procedure aims to control the FDR below α level by rejecting $\{H_{(1)}, \ldots, H_{(k)}\}$ where $k = \max_i \left\{p_{(i)} < \frac{\alpha i}{m}\right\}$.
- With the BH procedure, $FDR < \alpha \cdot \frac{m_0}{m}$.

2 FDR controlling confidence sets

2.1 Confidence Sets (two-sided)

In constructing the confidence sets for the area in the image $\{s \in S : \mu(s) > c\}$, we define $\hat{A}_c$ as $\hat{A}_c = \{s \in S : \hat{\mu}(s) \geq c\}$ which could be thought of as a point estimate for the area above c. Take

$$t(s) = \frac{\hat{\mu}(s) - c}{\hat{\sigma}(s)/\sqrt{n}}$$

the t-statistics for testing $\{H_0 : \mu(s) = c\}$. The FDR-controlling two-sided confidence sets constructed via testing are obtained by the following procedure:

1. Obtain the p-values $2 \cdot \{1 - T_{n-1}(|t(s)|)\}$ where T_{n-1} is the cdf of t-distribution with $n-1$ degrees of freedom.
2. Apply the Benjamini-Hochberg procedure to the p-values at α level to get the set of $s \in S$ rejected by the BH procedure i.e., $\mathcal{B} = \{s \in S : p_s \leq p_k\}$
3. Set the upper confidence set $\hat{\mathcal{A}}_c^+$ and the lower confidence set $\hat{\mathcal{A}}_c^-$ as
$$\hat{\mathcal{A}}_c^+ = \{\hat{\mathcal{A}}_c \cap \mathcal{B}\}$$
and
$$\hat{\mathcal{A}}_c^- = \{\hat{\mathcal{A}}_c \cup \mathcal{B}^C\}$$
.

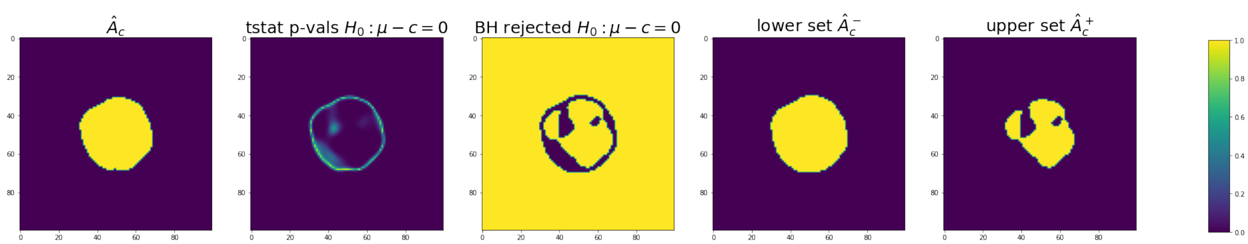

Figure 1: plot displaying the confidence set constructing procedures based on two-sided test representing $\hat{A}_c$, t-statistics p-values, $\mathcal{B}$, $\hat{A}_c^-$, $\hat{A}_c^+$ from left to right.

The error rate for the two-sided confidence sets is given as:

$$\mathbb{E}\left(\frac{|\hat{\mathcal{A}}_c^+ \setminus \mathcal{A}_c \cup \mathcal{A}_c \setminus \hat{\mathcal{A}}_c^-|}{|\hat{\mathcal{A}}_c^+ \cup (\hat{\mathcal{A}}_c^-)^C|}\right) \tag{1}$$

The confidence sets defined as above are expected to keep the error rate 1 below α level.

2.2 Confidence Sets (one-sided)

To construct the confidence set for the area $\{s \in S : \mu(s) \geq c\}$ with $\hat{A}_c = \{s \in S : \hat{\mu}(s) \geq c\}$, we can think of a testing approach using one-sided test. First, consider the null hypothesis $\{H_0 : \mu(s) < c\}$. Then, the FDR-controlling one-sided upper confidence set is obtained by the following procedure:

1. Obtain the p-values $\{1 - T_{n-1}(t(s))\}$ where T_{n-1} is the cdf of t-distribution with $n-1$ degrees of freedom.

2. Apply the Benjamini-Hochberg procedure to the p-values at α level to get the set of $s \in S$ rejected by the BH procedure i.e., $\mathcal{B} = \{s \in S : p_s \leq p_k\}$

3. Set the upper confidnce set $\hat{\mathcal{A}}_c^+$

$$\hat{\mathcal{A}}_c^+ = \{\hat{\mathcal{A}}_c \cap \mathcal{B}\}$$

.

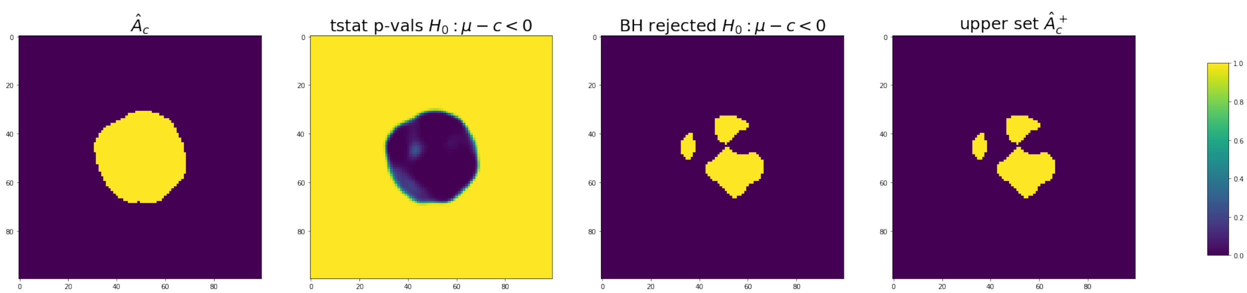

Figure 2: plot displaying the confidence set constructing procedures based on one-sided test representing $\hat{A}_c$, t-statistics p-values, $\mathcal{B}$, and $\hat{A}_c^+$ from left to right.

Conversely, to construct the confidence set for the area $\{s \in S : \mu(s) \leq c\}$ via one-sided test using the null hypothesis $\{H_0 : \mu(s) > c\}$ and $\hat{A}_c = \{s \in S : \hat{\mu}(s) \leq c\}$, we use the following procedure:

1. Obtain the p-values $\{T_{n-1}(t(s))\}$ where T_{n-1} is the cdf of t-distribution with $n-1$ degrees of freedom.

2. Apply the Benjamini-Hochberg procedure to the p-values at α level to get the set of $s \in S$ rejected by the BH procedure i.e., $\mathcal{B} = \{s \in S : p_s \leq p_k\}$

3. Set the lower confidnce set $\hat{\mathcal{A}}_c^-$

$$\hat{\mathcal{A}}_c^- = \{\hat{\mathcal{A}}_c \cup \mathcal{B}^C\}$$

.

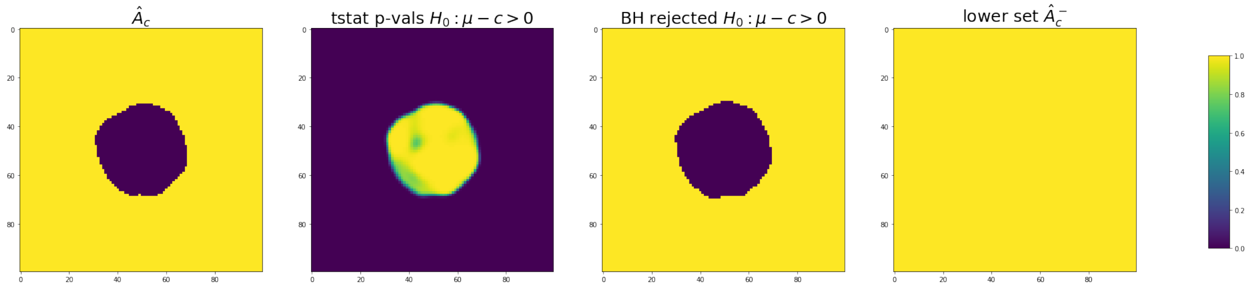

Figure 3: plot displaying the confidence set constructing procedures based on one-sided test representing $\hat{A}_c$, t-statistics p-values, $\mathcal{B}$, and $\hat{A}_c^-$ from left to right.

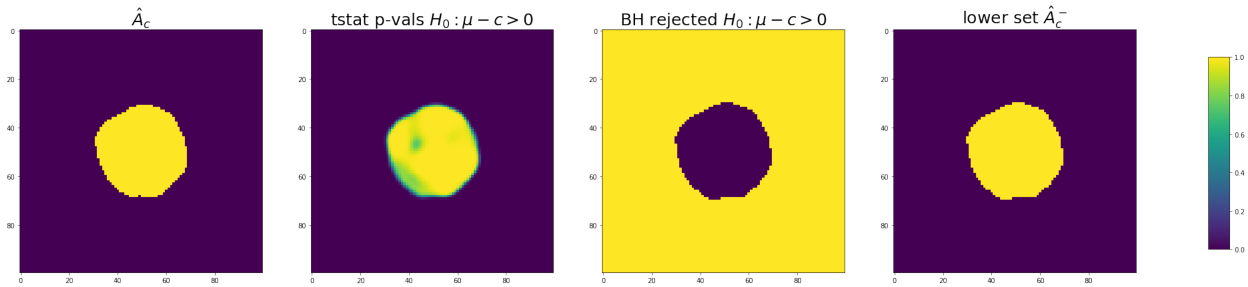

Figure 4: plot displaying the confidence set constructing procedures based on one-sided test representing $\hat{A}_c$, t-statistics p-values, $\mathcal{B}$, and $\hat{A}_c^-$ from left to right.

The error rate for one-sided upper and lower confidence sets are measured by the following respctively:

$$\mathbb{E}\left(\frac{|\hat{\mathcal{A}}_c^+ \setminus \mathcal{A}_c|}{|\hat{\mathcal{A}}_c^+|}\right) \tag{2}$$

$$\mathbb{E}\left(\frac{|\mathcal{A}_c \setminus \hat{\mathcal{A}}_c^-|}{|(\hat{\mathcal{A}}_c^-)^C|}\right) \tag{3}$$

! maybe to use the above error rate (3), we should use $\hat{A}_c = \{s \in S : \hat{\mu}(s) \geq c\}$ instead of $\hat{A}_c = \{s \in S : \hat{\mu}(s) \leq c\}$?? !

3 Simulation

3.1 Simulation Setting

Take $50 * 50$ and $100 * 100$ Gaussian random field with circular signal where $\mu = 3$:

The noise can be applied to the the above four to form a random field with different noise settings. For each voxels, the noise is drawn from $\mathcal{N}(0, 3)$. The smoothing was done with FWHM=8.

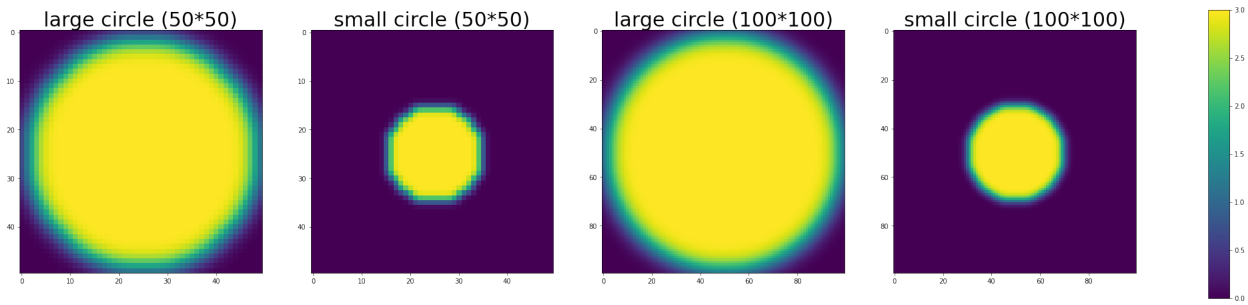

Figure 5: Four different signal settings: small or large circle for 50*50 and 100*100 image size.

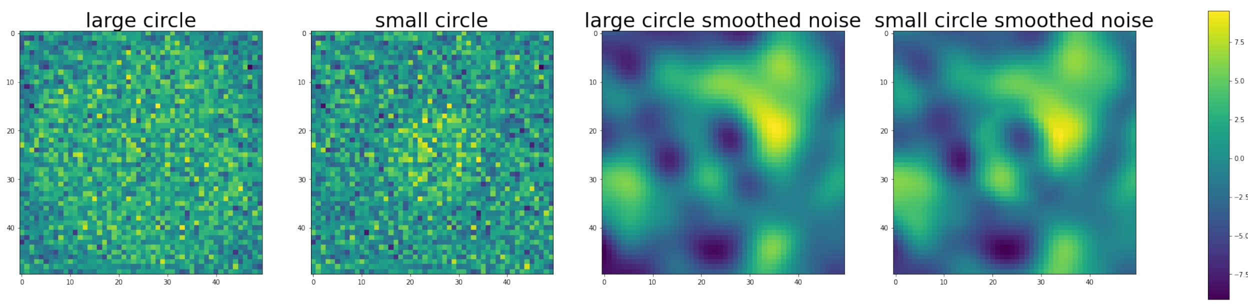

Figure 6: Two different noise settings: smoothed or uncorrelated noise.

3.2 Confidence Sets

Figure 7 shows the upper and lower confidence sets drawn on top of the signal. The confidence sets were constructed for 8 different images, each a combination of small/large circle, 100*100/50*50 image size, and smoothed/uncorrelated noise.

3.3 Error Rate

3.3.1 Two-sided

We look at the false discovery rate, expressed as (1), calculated from 1,000 simulations. Recall that $FDR < \alpha \cdot \frac{m_0}{m}$.

- Since on the discrete grid, there was often no voxel exactly of the value of the threshold $c = 0.5, 2, 3$. Instead, range was used to capture the number of voxels m_0. In this simulation, used range of $\{c \pm 0.2\}$
- For $c = 0.5$, since m_0 is 0, the error rates are very close to 0.
- For $c = 2$ and $c = 3$, the error rates are all below the $\alpha \frac{m_0}{m}$ level.
- Generally no big difference between image size 50*50 and 100*100.

Now we look at FWER which is $1 - \mathbb{P}[\hat{\mathcal{A}}_c^+ \subseteq \mathcal{A}_c \subseteq \hat{\mathcal{A}}_c^-]$.The FWER table is calculated from 1,000 simulations. Here the FWER is not controlled under $\alpha = 0.05$ level as he exclusion rate is well above 0.05. The uncorrelated noise images have higher exclusion rate due to the confidence sets having more erratic shapes in such cases.

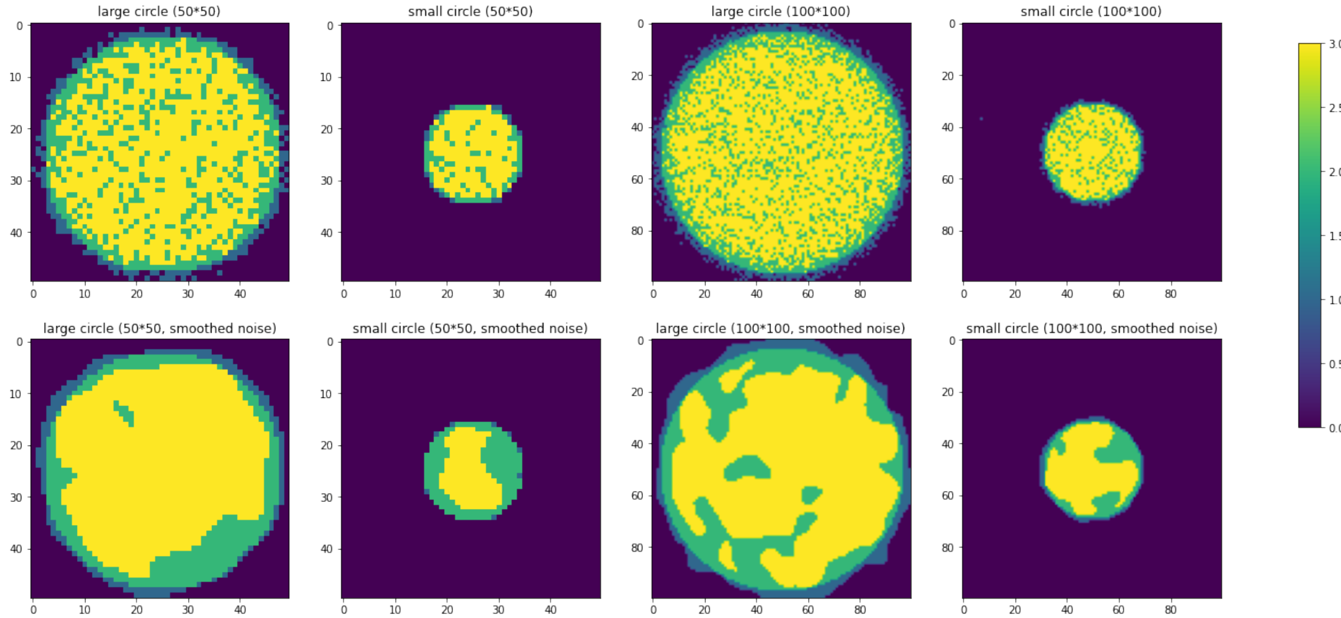

Figure 7: Confidence sets and signal

3.3.2 One-sided

	small	**small&smth**	$\alpha\frac{m_0}{m}$	**large**	**large&smth**	$\alpha\frac{m_0}{m}$
50*50						
c=0.5	0.0002	0.0002	**0.0**	0.0009	0.0009	**0.0**
c=0.5/std=*3	0.0041	0.0021		0.0026	0.0029	
c=2	0.0002	0.0002	**0.00072**	0.0006	0.0006	**0.00168**
c=2/std=*3	0.0005	0.0006		0.0017	0.0014	
c=3	0.0021	0.0024	**0.00466**	0.012	0.0117	**0.02674**
c=3/std=*3	0.0024	0.0022		0.0135	0.0122	
100*100						
c=0.5	0.0003	0.0003	**0.0**	0.0007	0.0007	**0.0**
c=0.5/std=*3	0.0038	0.0053		0.0023	0.0024	
c=2	0.0001	0.0001	**0.00038**	0.0007	0.0007	**0.00238**
c=2/std=*3	0.0004	0.0004		0.002	0.0015	
c=3	0.0022	0.0023	**0.00461**	0.013	0.0131	**0.02817**
c=3/std=*3	0.0024	0.0026		0.0142	0.0133	

Table 1: FDR

	small	**small&smth**	**large**	**large&smth**
50*50				
c=0.5	0.131	0.055	0.809	0.446
c=0.5/std=*3	0.189	0.032	0.89	0.354
c=2	0.364	0.196	0.697	0.352
c=2/std=*3	0.295	0.101	0.086	0.049
c=3	0.992	0.268	1	0.483
c=3/std=*3	0.985	0.267	0.984	0.266
100*100				
c=0.5	0.558	0.238	0.998	0.712
c=0.5/std=*3	0.581	0.102	1	0.409
c=2	0.656	0.372	0.996	0.743
c=2/std=*3	0.674	0.199	0.224	0.075
c=3	1	0.666	1	0.915
c=3/std=*3	1	0.635	1	0.611

Table 2: $FWER$