# Confidence Sets Controlling for False Discovery Rate

Howon Ryu

March 6, 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 $\alpha$ 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 following procedure describes the construction of the FDR-controlling two-sided confidence sets via the Benjamini-Hochberg 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 $\alpha$ 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
$$\hat{\mathcal{A}}_c^- = \{\hat{\mathcal{A}}_c \cup \mathcal{B}^C\}$$
.

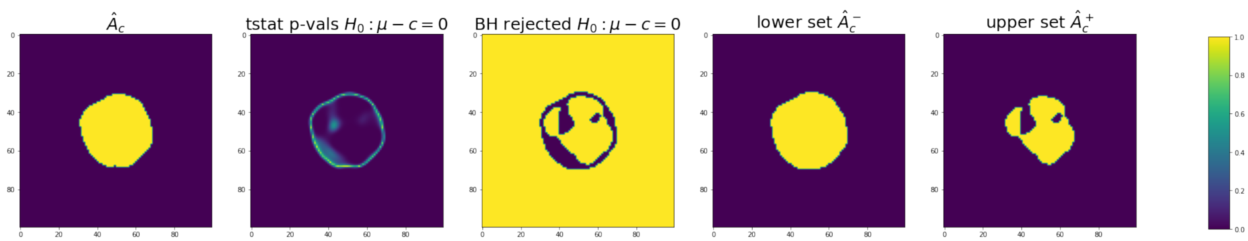


Figure 1: plot displaying the confidence set constructing procedures based on two-sided test representing $\hat{A}_c$, t-statistics p-values, $\mathcal{B}$, $\hat{A}_c^-$, $\hat{A}_c^+$ from left to right.

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 $\alpha$ 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 $\alpha$ 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 $\hat{\mathcal{A}}_c^+$

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

.

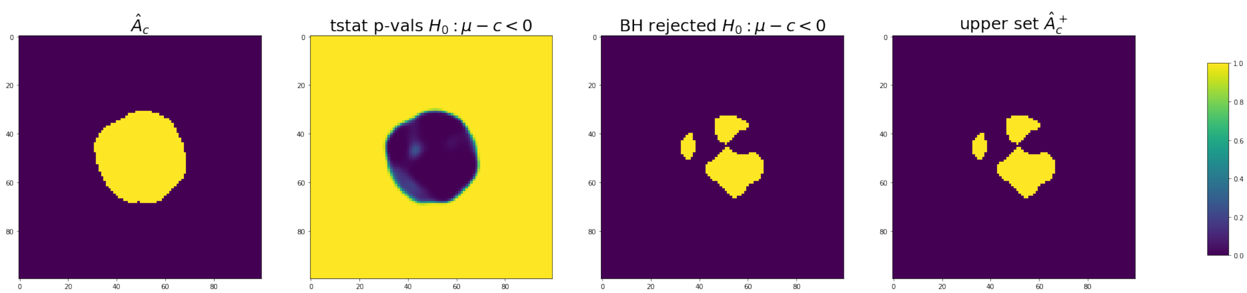


Figure 2: plot displaying the confidence set constructing procedures based on one-sided test representing $\hat{A}_c$, t-statistics p-values, $\mathcal{B}$, and $\hat{A}_c^+$ from left to right.

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 $\alpha$ 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 $\hat{\mathcal{A}}_c^-$

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

.

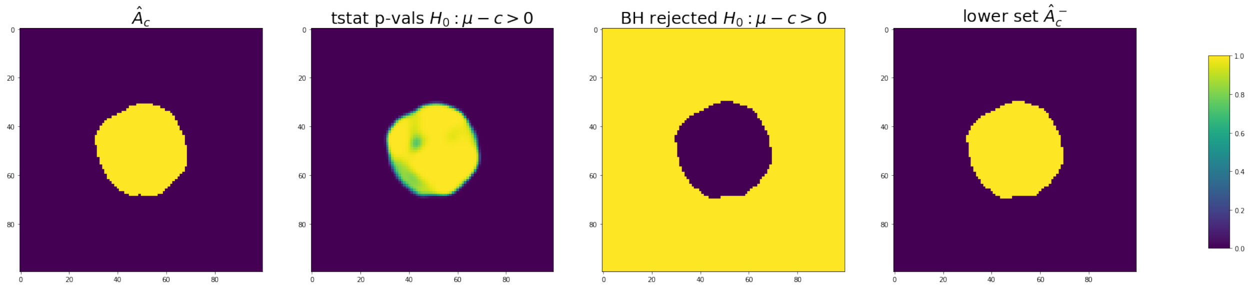


Figure 3: plot displaying the confidence set constructing procedures based on one-sided test representing $\hat{A}_c$, t-statistics p-values, $\mathcal{B}$, and $\hat{A}_c^-$ from left to right.

The error rate for one-sided upper sets is measured by

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

and the lower set is measured by

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

### 2.3 Power and false non-discovery rate

## 3 Adaptive procedure

As the Benjamini-Hochberg procedure controls the FDR at $\alpha \cdot \frac{m_0}{m}$ level in actuality, this often leads to conservative inference on the confidence sets. Adaptive procedures remedy this by using the threshold collection of the form $\Delta_i = \pi_0^{-1}\frac{\beta(i)}{m}\alpha$ where $\beta(i)$ is the shape function for a step-up procedure. By multiplying the threshold by a factor of $\pi_0^{-1} = \frac{m}{m_0}$, we expect to have less conservative procedure. In practice, we use $\Delta_i = \hat{\pi}_0^{-1}\frac{\beta(i)}{m}\alpha$ where $\hat{\pi}_0^{-1} = \frac{m}{\hat{m}_0}$ is the estimator of the true $\pi_0^{-1}$.

### 3.1 Two-stage adaptive procedure

Blanchard and Roquain [1] suggests a two-stage adaptive procedure that provably controls the FDR at $\alpha$ level. In the first stage, a step-up procedure is conducted with threshold collection $\Delta_i = \frac{\beta(i)}{m}\alpha_0$ to get an estimate $\hat{m}_0 = m - |R_0|$ where $|R_0|$ is the number of rejected hypotheses from the first stage. In the second stage, an adaptive procedure is conducted with $\Delta_i = \hat{\pi}_0^{-1}\frac{\beta(i)}{m}\alpha_1$ where $\hat{\pi}_0^{-1} = F_\kappa(\frac{\hat{m}_0}{m})$. $F_\kappa(\cdot)$ is given as follows:

$$F_\kappa(x) = \begin{cases} 1, & \text{if } x \leq \kappa^{-1} \\ \frac{2\kappa^{-1}}{1-\sqrt{1-4(1-x)\kappa^{-1}}}, & \text{otherwise} \end{cases}$$

Blanchard and Roquain states that this two-stage procedure with $\alpha_0 = \alpha/4, \alpha_1 = \alpha/2$ and $\kappa = 2$ is less conservative compared to the linear step-up procedure alone (Benjamini-Hochberg) when the large proportion of the rejection is expected, or potentially when $\hat{\pi_0}$ is small.

# 4 Simulation

## 4.1 Simulation Setting

We now construct confidence sets based on different signal shapes, image sizes, and noise settings. Take circular and ramp signal of image sizes $50 * 50$ and $100 * 100$ as presented in figure 4. Then, correlated(smoothed) or uncorrelated noise fields are added to the the above six signals to form a random field with.

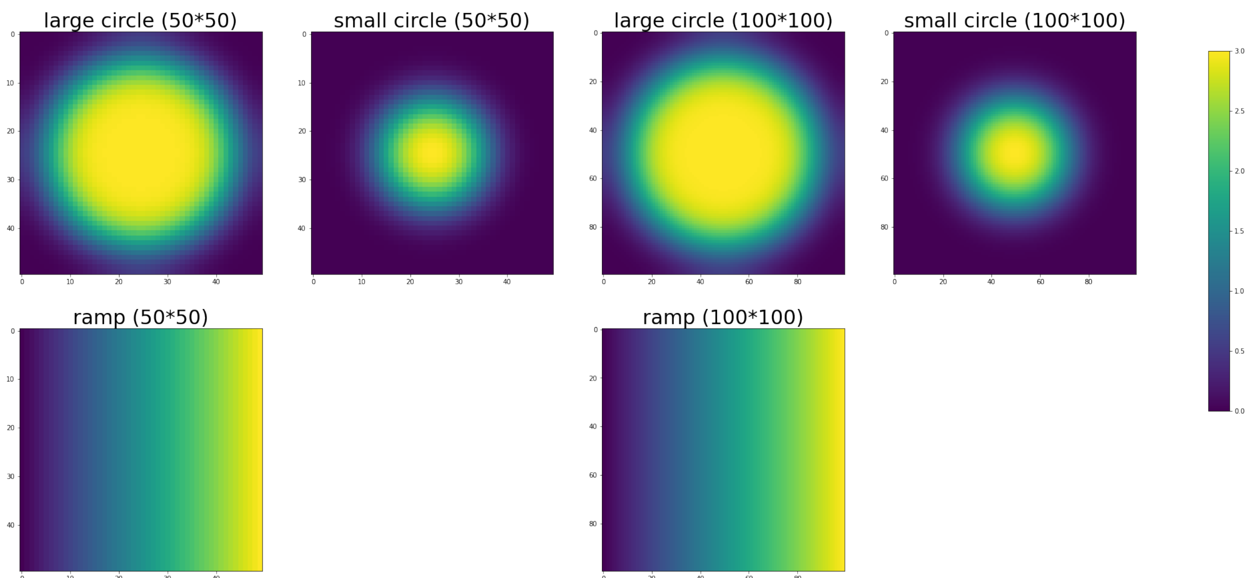


Figure 4: six different signal settings: small or large circle for 50*50 and 100*100 image size, and ramp for 50*50 and 100*100 image size

## 4.2 Confidence Sets

Figure 5 and 6 shows the upper and lower confidence sets superimposed on the signal. The confidence sets were constructed for six different images, each a combination of different signals (small circle/large circle/ramp) and noises (correlated/uncorrelated).

The confidence sets constructed by the two-stage adaptive procedure shows bigger $\hat{A}_c^+$ and $\hat{A}_c+^{-c}$ than the ones constructed by the Benjamini-Hochberg procedure which suggests that the two-stage procedure is less conservative.

——————————-WORK FROM ERE——————————-

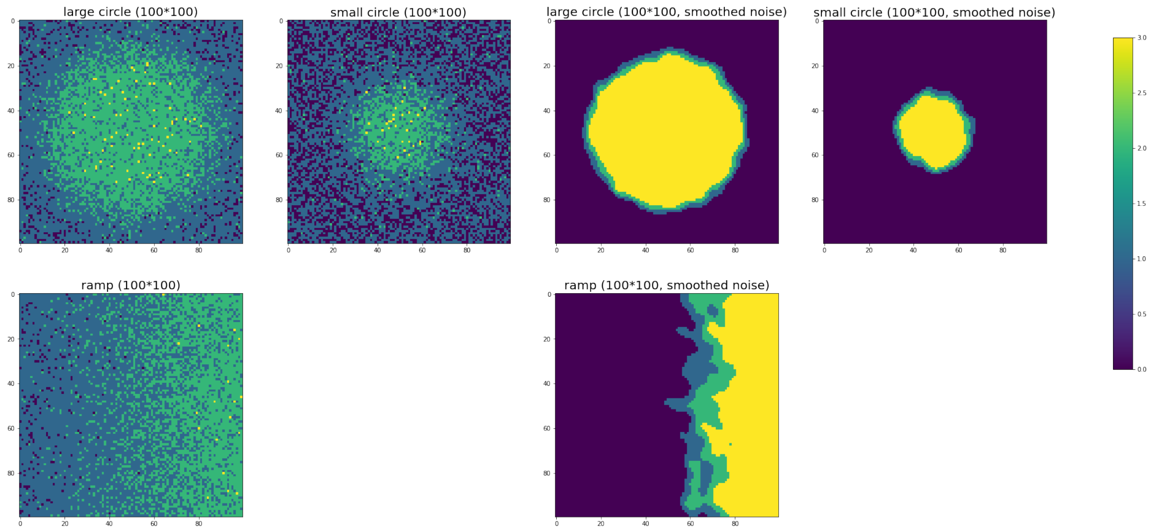


Figure 5: $\hat{A}_c^+ + \hat{A}_c + \hat{A}_c^-$ by the Benjamini-Hochberg procedure. Yellow area represents $A_c^+$, green area including the yellow area represents $\hat{A}_c$, and the blue area including green and yellow area represents $\hat{A}_c^-$.

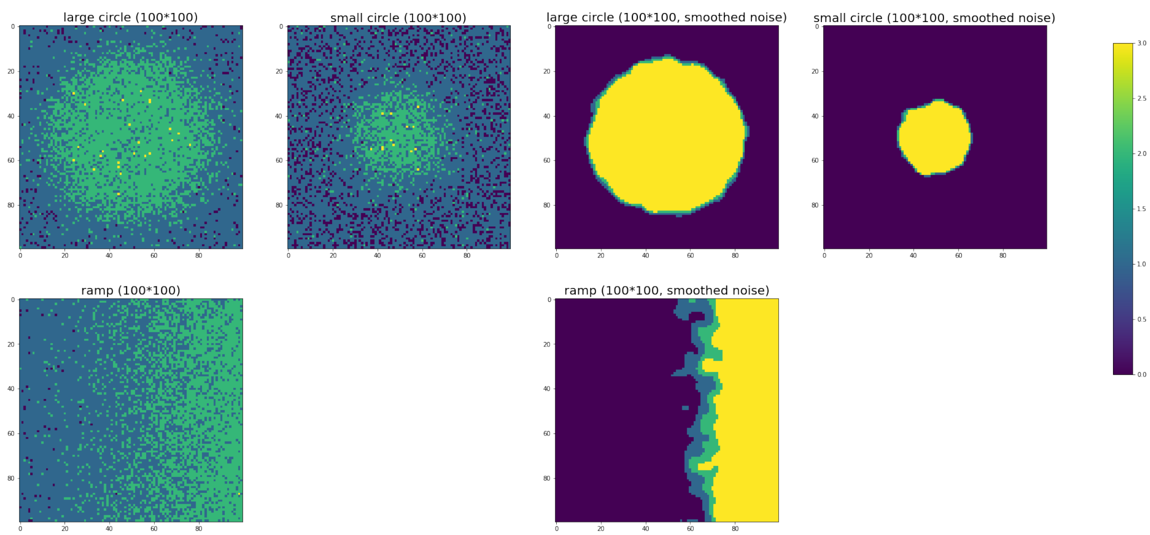


Figure 6: $\hat{A}_c^+ + \hat{A}_c + \hat{A}_c^-$ by two-stage adaptive procedure. Yellow area represents $A_c^+$, green area including the yellow area represents $\hat{A}_c$, and the blue area including green and yellow area represents $\hat{A}_c^-$.

## 4.3 Error Rate

### 4.3.1 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 $\{c \pm 0.2\}$

| | **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: $FDR$

- 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.

# References


[1] Gilles Blanchard and Etienne Roquain. "Adaptive False Discovery Rate Control under Independence and Dependence." In: *Journal of Machine Learning Research* 10.12 (2009).