A Statistical Comparison of Call Volume Uniformity Due to Mailing Strategy

Andrew M. Raim

Center for Statistical Research and Methodology
U.S. Census Bureau
andrew.raim@census.gov

2018 Joint Statistical Meetings
Vancouver, CA

Joint work with **Elizabeth Nichols** (CES, U.S. Census Bureau) and **Thomas Mathew** (CSRM, U.S. Census Bureau)

Introduction

- The Census Bureau will send mail to each household in the U.S. to request participation in the 2020 Census.
- Responses will be encouraged on a large scale through the internet and by phone (U.S. Census Bureau, 2017).
- Telephone helplines will be highlighted in mailings, both to assist with internet responses and to serve as a mode of response itself.
- Ideally, the Census Bureau would prefer uniform volume of calls throughout each week of the census. This would require fewer helpline staff to cover peak times, and a more constant workload for hired staff.

Introduction

Scheduling Mailings

- The Census Bureau determines the schedule of mailing materials, which influences when call volumes tend to occur (Chesnut, 2003; Zajac, 2012).
- Nichols et al. (2018) notes some general patterns in call volumes.
    1. Peaks occur on the expected mail delivery dates.
    2. Volumes are highest on Mondays and Tuesdays, decline through the rest of the week, and are lowest weekends.
    3. Volumes diminish after Census Day.
    4. Volumes diminish after all mailings have occurred.
- The Census Bureau is considering plans to split recipients into two or more groups and stagger mailings to arrive on different days of the week.

Introduction

Datasets

- We compare unifomity among call volumes recorded in three census experiments. These operations are referred to as National Census Bureau Surveys (NCBS's) in mailing materials.
- An unstaggered mailing strategy was used in the **2016 September NCBS** (Eggleston and Coombs, 2017) and **2016 June NCBS** (Coombs, 2017).
- A staggered mailing schedule was used in the **2017 March NCBS** (Nichols et al., 2018). Here, study participants were randomly assigned into either a Monday Mailout group or a Thursday Mailout group.
- Live agents were not present to answer the helpline and callers received a prerecorded message.
- Caller identities were not recorded, so we cannot distinguish whether multiple calls were made by the same caller.

Mailing Schedules

(a) The 2017 March NCBS.

	Monday Mailout	Thursday Mailout
1	Mon 3/06/2017	Thu 3/09/2017
2	Thu 3/09/2017	Mon 3/13/2017
3	Mon 3/20/2017	Thu 3/23/2017
4	Mon 3/27/2017	Thu 3/30/2017

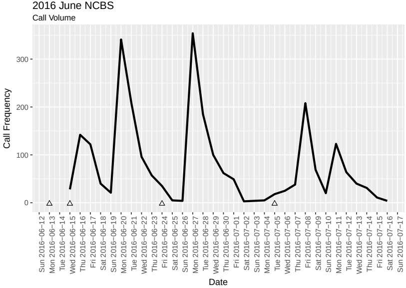
(b) 2016 Sept NCBS.

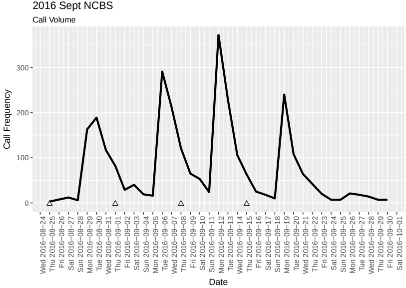
	Date
1	Thu 8/25/2016
2	Thu 9/01/2016
3	Thu 9/08/2016
4	Thu 9/15/2016

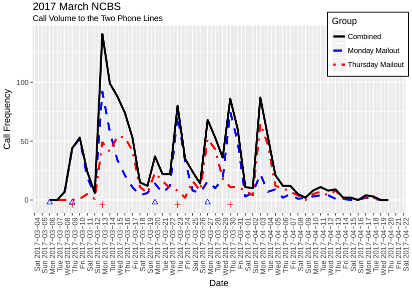
(c) 2016 June NCBS.

	Date
1	Mon 6/13/2016
2	Wed 6/15/2016
3	Fri 6/24/2016
4	Tue 7/05/2016

Observed Call Data







Inference Approach

- To our knowledge, inference comparing the unifomity of two discrete distributions is not standard. We consider Kullback-Leibler (K-L) divergence and entropy and make use of basic large sample theory.
- Many conventional tests are designed to detect departure from equality; e.g. chi-square and Kolmogorov-Smirnov tests.
- Cover and Thomas (2006) introduces K-L distance, entropy, and related concepts, and discusses fundamental applications in information theory.
- K-L divergence and entropy have been used to justify information criteria (Konishi and Kitagawa, 2008), to obtain variational approximations to complicated distributions (Ormerod and Wand, 2010; Blei et al., 2017), and as a basis for inference (Pardo, 2006; Girardin and Lequesne, 2017).
- Paninski (2008) tests for departure between single discrete distribution and discrete uniform in a sparse setting (many categories and few observations).

Quantifying "More Uniform"

- Suppose $\boldsymbol{p} = (p_1, \dots, p_k)$ and $\boldsymbol{q} = (q_1, \dots, q_k)$ are probability distributions on categories labeled $(1, \dots, k)$.
- Let $D(\boldsymbol{p}, \boldsymbol{q}) = \sum_{j=1}^k p_j \log(p_j / q_j)$ be the Kullback-Leibler (K-L) divergence.
- Let $\bar{\boldsymbol{e}} = (1/k \dots, 1/k)$ denote the discrete uniform distribution and $\boldsymbol{e}_j$ denote a vector with 1 in the jth entry and zeros elsewhere.
- We will say that $\boldsymbol{q}$ is a "more uniform" distribution than $\boldsymbol{p}$ if

$$D(\boldsymbol{p}, \bar{\boldsymbol{e}}) > D(\boldsymbol{q}, \bar{\boldsymbol{e}}) \quad \iff \quad \mathcal{E}(\boldsymbol{p}) < \mathcal{E}(\boldsymbol{q}),$$

  where $\mathcal{E}(\boldsymbol{p}) = -\sum_{j=1}^k p_j \log p_j$ is the entropy.
- The entropy of any $\boldsymbol{p}$ is bounded, with

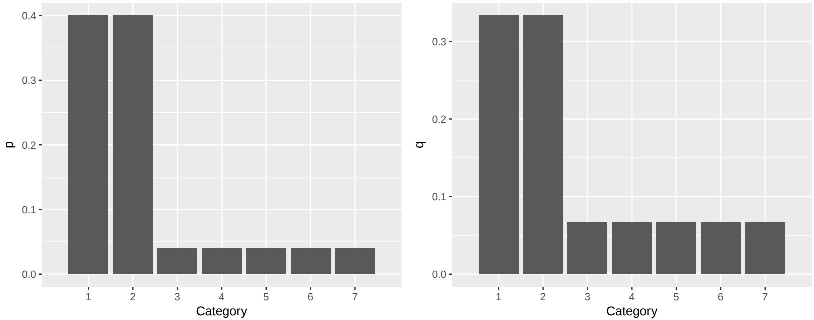
$$\begin{aligned} &\mathcal{E}(\boldsymbol{p}) \leq \mathcal{E}(\bar{\boldsymbol{e}}) = \log k, \\ &\mathcal{E}(\boldsymbol{p}) \geq \mathcal{E}(\boldsymbol{e}_j) = 0, \quad \text{for any } j = 1, \dots, k. \end{aligned}$$

Quantifying “More Uniform”

- Suppose $\boldsymbol{p}$ and $\boldsymbol{q}$ are parameterized by $\boldsymbol{\theta}$ (which depends on the choice of model).
- Our quantity of interest is the difference in entropy $g(\boldsymbol{\theta}) = \mathcal{E}(\boldsymbol{q}) - \mathcal{E}(\boldsymbol{p})$.
- This quantity is bounded, with $-\log k \leq g(\boldsymbol{\theta}) \leq \log k$.
- We will consider testing hypotheses of the form

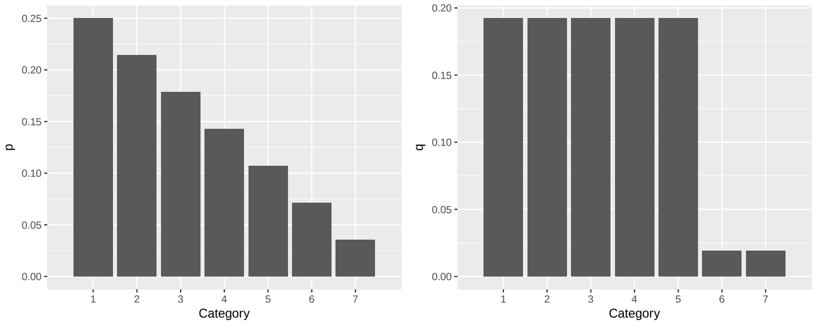
$$\begin{aligned} H_0 &: g(\boldsymbol{\theta}) = 0 \quad \text{vs.} \quad H_1 : g(\boldsymbol{\theta}) \neq 0, \\ H_0 &: g(\boldsymbol{\theta}) \leq 0 \quad \text{vs.} \quad H_1 : g(\boldsymbol{\theta}) > 0, \\ H_0 &: g(\boldsymbol{\theta}) \geq 0 \quad \text{vs.} \quad H_1 : g(\boldsymbol{\theta}) < 0. \end{aligned}$$

Example 1



- Dist'ns: $\boldsymbol{p} = \frac{1}{25}(10, 10, 1, 1, 1, 1, 1)$ and $\boldsymbol{q} = \frac{1}{15}(5, 5, 1, 1, 1, 1, 1)$.
- Entropies: $\mathcal{E}(\boldsymbol{p}) = 1.3768$ and $\mathcal{E}(\boldsymbol{q}) = 1.6351$.
- Entropy difference: $g(\boldsymbol{\theta}) = \mathcal{E}(\boldsymbol{q}) - \mathcal{E}(\boldsymbol{p}) = 0.2583$.

Example 2



- Dist'ns: $\boldsymbol{p} = \frac{1}{28}(7, 6, 5, 4, 3, 2, 1)$ and $\boldsymbol{q} = \frac{1}{15}(10, 10, 10, 10, 10, 1, 1)$.
- Entropies: $\mathcal{E}(\boldsymbol{p}) = 1.8091$ and $\mathcal{E}(\boldsymbol{q}) = 1.7372$.
- Entropy difference: $g(\boldsymbol{\theta}) = \mathcal{E}(\boldsymbol{q}) - \mathcal{E}(\boldsymbol{p}) = -0.0719$.

Multinomial Model

- We want to compare two census experiments where I mailing schedules were used among them.
- For each mailing schedule, calls were collected for J weeks.
- Let $\boldsymbol{X}_{ij} = (X_{ij1}, \ldots, X_{ijk})$ be the call counts observed on (Sun, Mon, ..., Sat) for the jth week and the ith mailing schedule, $i = 1, \ldots, I$ and $j = 1, \ldots, J$.
- Assume that

$$\boldsymbol{X}_{ij} \stackrel{\text{ind}}{\sim} \text{Mult}_k(m_{ij}, \boldsymbol{p}_{ij}), \quad \text{where } \boldsymbol{p}_{ij} = (p_{ij1}, \ldots, p_{ijk})$$

  is the (unknown) day-of-week distribution and $m_{ij} = \sum_{\ell=1}^{k} X_{ij\ell}$ is the (fixed) total call count.
- Let $\boldsymbol{\theta} = (\boldsymbol{p}_{11}, \ldots, \boldsymbol{p}_{IJ})$ be the unknown probabilities and $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{p}}_{11}, \ldots, \hat{\boldsymbol{p}}_{IJ})$ be the sample proportions with $\hat{\boldsymbol{p}}_{ij} = \boldsymbol{X}_{ij}/m_{ij}$.

Scenario S1

- This scenario consists of two census experiments with one mailing schedule used in each.
- This covers the "2016 June NCBS vs. 2016 Sept NCBS" analysis.
- Here we are interested in

$$g_j(\boldsymbol{\theta}) = \mathcal{E}(\boldsymbol{p}_{2j}) - \mathcal{E}(\boldsymbol{p}_{1j}),$$

for each week $j = 1, \dots, J$.

Scenario S2

- This scenario consists of two census experiments with one mailing schedule used in the first and two used in the second.
- This covers the "2016 Sept NCBS versus 2017 March NCBS" and "2016 June NCBS versus 2017 March NCBS" analyses.
- Let $\boldsymbol{q}_j = (q_{j1}, \dots, q_{jk})$ be the overall day-of-week distribution for calls from the jth week of the second experiment.
- Let π_j be the probability of a call received during week j being from first mailing schedule.
- By the law of total probability,

$$\begin{aligned} q_{j\ell} &= \text{P}\{\text{Call occurs on day-of-week } \ell \mid \text{Schedule 1}\}\,\text{P}\{\text{Schedule 1}\} \\ &\quad + \text{P}\{\text{Call occurs on day-of-week } \ell \mid \text{Schedule 2}\}\,\text{P}\{\text{Schedule 2}\} \\ &= \pi_j p_{2j\ell} + (1 - \pi_j) p_{3j\ell}. \end{aligned}$$

Then we may write $\boldsymbol{q}_j = \pi_j \boldsymbol{p}_{2j} + (1 - \pi_j)\boldsymbol{p}_{3j}$, and our ultimate quantities of interest are

$$g_j(\boldsymbol{\theta}) = \mathcal{E}(\boldsymbol{q}_j) - \mathcal{E}(\boldsymbol{p}_{1j}), \quad j = 1, \dots, J.$$

Designating Weeks

- For each census experiment, we designate day 1 as the day of the first mailing. For the 2017 March NCBS, day 1 is the Monday of the very first mailing.
- We then designate week 1 as days 1–7, week 2 as days 8–14, and so on.
- Weekends were kept intact rather than being combined or discarded.
- We consider weeks 1–5 in each census experiment, and disregard calls which occurred in week 6 or later because call activity became sparse.
- According to our definition of weeks, Thursday Mailout group calls for the 2017 March NCBS will be extremely unlikely. Therefore, staggering has very little effect on week 1.
- Alternative definitions for weeks can change the results.

Tests and Confidence Limits

- Let $\boldsymbol{c} = (c_1, \ldots, c_I)$ and $\boldsymbol{d} = (d_1, \ldots, d_I)$ be given probability distributions on $\{1, \ldots, I\}$.
- Let us generally write for the jth week,

$$g_j(\boldsymbol{\theta}) = \mathcal{E}(c_1 \boldsymbol{p}_{1j} + \cdots + c_I \boldsymbol{p}_{Ij}) - \mathcal{E}(d_1 \boldsymbol{p}_{1j} + \cdots + d_I \boldsymbol{p}_{Ij}).$$

- We do not have a situation where two census experiments used data from a common mailing schedule. Therefore, in practice, we will have $c_i d_i = 0$ for $i = 1, \ldots, I$.

Tests and Confidence Limits

- For large samples (large m_{ij}), we approximately have

$$\hat{\boldsymbol{\theta}} \sim \mathrm{N}(\boldsymbol{\theta}, \boldsymbol{\Sigma}),$$

$$\boldsymbol{\Sigma} = \text{Blockdiag}\left(m_{ij}^{-1}\left[\text{Diag}(\boldsymbol{p}_{ij}) - \boldsymbol{p}_{ij}\boldsymbol{p}_{ij}^\top\right] : i = 1, \ldots, I; j = 1, \ldots, J\right).$$

- By the Delta method, we approximately have (for large m_{ij})

$$g_j(\hat{\boldsymbol{\theta}}) \sim \mathrm{N}\left(g_j(\boldsymbol{\theta}), \sigma^2_{g_j(\boldsymbol{\theta})}\right), \quad \text{where}$$

$$\sigma^2_{g_j(\boldsymbol{\theta})} = \left[\frac{\partial g_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]^\top \boldsymbol{\Sigma} \left[\frac{\partial g_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right],$$

$$\frac{\partial g_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \boldsymbol{e}_j \otimes \left[c_1 \nabla \mathcal{E}\left(\sum_{i=1}^I c_i \boldsymbol{p}_{ij}\right) - d_1 \nabla \mathcal{E}\left(\sum_{i=1}^I d_i \boldsymbol{p}_{ij}\right)\right] \\ \vdots \\ \boldsymbol{e}_j \otimes \left[c_I \nabla \mathcal{E}\left(\sum_{i=1}^I c_i \boldsymbol{p}_{ij}\right) - d_I \nabla \mathcal{E}\left(\sum_{i=1}^I d_i \boldsymbol{p}_{ij}\right)\right] \end{pmatrix},$$

$$\nabla \mathcal{E}(\boldsymbol{p}) = (-\log(p_2/p_1), \ldots, -\log(p_k/p_1))^\top.$$

- Here, $\otimes$ denotes the matrix Kronecker product and the first probability $p_{ij1} = 1 - \sum_{\ell=2}^k p_{ij\ell}$ is taken to be the baseline for each i, j.

United States Census Bureau

Tests and Confidence Limits

- Let $\mathcal{Z} = g_j(\hat{\boldsymbol{\theta}}) \Big/ \sqrt{\sigma^2_{g_j(\hat{\boldsymbol{\theta}})}}$ where $g_j(\hat{\boldsymbol{\theta}}) \dot{\sim} \text{N}\left(g_j(\boldsymbol{\theta}), \sigma^2_{g_j(\boldsymbol{\theta})}\right)$.

- Under the restriction $g_j(\boldsymbol{\theta}) = 0$, we have $\mathcal{Z} \dot{\sim} \text{N}(0, 1)$.

- $\mathcal{Z}$ tests corresponding to our stated three hypotheses are
  1. Reject $H_0 : g(\boldsymbol{\theta}) = 0$ if $|\mathcal{Z}| > z_{\alpha/2}$,
  2. Reject $H_0 : g(\boldsymbol{\theta}) \leq 0$ if $\mathcal{Z} > z_{\alpha}$,
  3. Reject $H_0 : g(\boldsymbol{\theta}) \geq 0$ if $\mathcal{Z} < z_{\alpha}$,

  where z_α are critical $\text{N}(0, 1)$ values.

- Corresponding $1 - \alpha$ level confidence limits are
  1. $g(\hat{\boldsymbol{\theta}}) \pm z_{\alpha/2}\sigma_{g(\hat{\boldsymbol{\theta}})}$,
  2. $g(\hat{\boldsymbol{\theta}}) - z_{\alpha}\sigma_{g(\hat{\boldsymbol{\theta}})}$,
  3. $g(\hat{\boldsymbol{\theta}}) + z_{\alpha}\sigma_{g(\hat{\boldsymbol{\theta}})}$.

2016 Sept NCBS vs. 2017 March NCBS

$$H_0 : g_j(\boldsymbol{\theta}) \leq 0 \quad \text{vs.} \quad H_1 : g_j(\boldsymbol{\theta}) > 0$$

H_0: "The day-of-week distribution in week j resulting from the 2016 September NCBS mailing schedule has ***larger or equal*** *entropy than the day-of-week distribution resulting from the 2017 March NCBS mailing schedule" versus H_1: "Not".*

Week	Est	SE	$\mathcal{Z}$-stat	p-value	CI Lo
1	0.0823	0.0602	1.3681	0.0856	0.0052
2	0.2605	0.0402	6.4731	4.802e-11	0.2090
3	0.1480	0.0411	3.6026	0.0002	0.0953
4	0.2273	0.0453	5.0166	2.629e-07	0.1693
5	-0.3376	0.0775	-4.3563	1.0000	-0.4369

For $\alpha = 0.10$, reject H_0 if $\mathcal{Z} > 1.282$.

2016 June NCBS vs. 2017 March NCBS

$$H_0 : g_j(\boldsymbol{\theta}) \leq 0 \quad \text{vs.} \quad H_1 : g_j(\boldsymbol{\theta}) > 0$$

H_0: "The day-of-week distribution in week j resulting from the 2016 June NCBS mailing schedule has ***larger or equal*** *entropy than the day-of-week distribution resulting from the 2017 March NCBS mailing schedule" versus H_1: "Not".*

Week	Est	SE	$\mathcal{Z}$-stat	p-value	CI Lo
1	-0.0153	0.0631	-0.2426	0.5958	-0.0962
2	0.3463	0.0379	9.1467	2.935e-20	0.2977
3	0.3905	0.0442	8.8356	4.980e-19	0.3338
4	0.3523	0.0565	6.2409	2.175e-10	0.2800
5	0.0253	0.0640	0.3956	0.3462	-0.0567

For $\alpha = 0.10$, reject H_0 if $\mathcal{Z} > 1.282$.

United States Census Bureau

2016 June NCBS vs. 2016 Sept NCBS

$$H_0 : g_j(\boldsymbol{\theta}) = 0 \quad \text{vs.} \quad H_1 : g_j(\boldsymbol{\theta}) \neq 0$$

*H_0: "The day-of-week distribution in week j resulting from the 2016 June NCBS mailing schedule has **equal** entropy to the day-of-week distribution resulting from the 2016 September NCBS mailing schedule" versus H_1: "Not".*

Week	Est	SE	$\mathcal{Z}$-stat	p-value	CI Lo	CI Hi
1	-0.0976	0.0451	-2.1630	0.0305	-0.1719	-0.0234
2	0.0857	0.0433	1.9782	0.0479	0.0144	0.1570
3	0.2425	0.0365	6.6499	2.934e-11	0.1825	0.3025
4	0.1250	0.0607	2.0602	0.0394	0.0252	0.2247
5	0.3629	0.0535	6.7823	1.183e-11	0.2749	0.4509

For $\alpha = 0.10$, reject H_0 if $|\mathcal{Z}| > 1.645$.

Conclusions

- We compared uniformity of two discrete distributions using procedures based on the Delta method.
- The staggered strategy yielded higher entropy than the two unstaggered experiments toward the middle of the study period, after both Monday and Thursday Mailout groups received their first mailing.
- However, the two unstaggered strategies also yielded different entropies from each other. This suggests that other factors beside staggering affect the uniformity.
- Entropy difference becomes smaller when component distributions are both closer to uniform, which makes differences harder to detect with the $\mathcal{Z}$-statistic.
- The Delta method estimator for entropy can be quite biased for small samples. It may be possible to improve small sample properties.

Contact: `andrew.raim@census.gov`

References I

David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112 (518):859–877, 2017.

John Chesnut. Telephone questionnaire assistance. Census 2000 Evaluation A.1.a, U.S. Census Bureau, 2003. URL `https://www.census.gov/pred/www/rpts/A.1.a.pdf`.

Julia Coombs. Analysis report for the small-scale mailout testing program June 2016 test on the placement and length of the user ID for an online Census Bureau survey. In *2020 Research and Testing*. 2017. (Internal Report).

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2nd edition, 2006.

Casey Eggleston and Julia Coombs. Effect of data use statements and postcard format on login rates for a mandatory online Census Bureau survey. In *2020 Research and Testing*. 2017. (Internal Report).

Valérie Girardin and Justine Lequesne. Entropy-based goodness-of-fit tests—a unifying framework: Application to DNA replication. *Communications in Statistics - Theory and Methods*, 2017. doi: 10.1080/03610926.2017.1401084.

References II

Sadanori Konishi and Genshiro Kitagawa. *Information criteria and statistical modeling*. Springer, 2008.

Elizabeth Nichols, Sarah Konya, Rachel Horwitz, and Andrew Raim. The effect of the mail delivery date on survey login rates and helpline call rates. In *2020 Research and Testing Report*. 2018. (Forthcoming).

J. T. Ormerod and M. P. Wand. Explaining variational approximations. *The American Statistician*, 64(2):140–153, 2010.

Liam Paninski. A coincidence-based test for uniformity given very sparsely sampled discrete data. *IEEE Transactions on Information Theory*, 54(10):4750–4755, 2008.

Leandro Pardo. *Statistical inference based on divergence measures*. Chapman & Hall/CRC, 2006.

U.S. Census Bureau. 2020 Census Operational Plan, 2017. URL `https://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/planning-docs/operational-plan.html`. Version 3.0.

Kevin Zajac. 2010 census telephone questionnaire assistance assessment report. 2010 Census Planning Memoranda Series No. 231, U.S. Census Bureau, 2012. URL `https://www.census.gov/2010census/pdf/2010_Census_TQA_Assessment.pdf`.

Call Volumes by Week

			2017 March		
Week	2016 June	2016 Sept	Mon Mailout	Th Mailout	Total
1	332	15	122	6	128
2	764	626	225	251	476
3	757	777	151	81	232
4	367	837	177	151	328
5	293	484	46	138	184
Total	2513	2739	721	627	1348

Weekly Estimates

2016 September NCBS vs. 2017 March NCBS

Week	$\hat{\boldsymbol{p}}(\texttt{Week})$							$\hat{\mathcal{E}}$
1	0.3857	0.3327	0.2388	0.0245	0.0122	0.0061	0.0000	1.2515
2	0.4224	0.3077	0.1190	0.0581	0.0421	0.0276	0.0232	1.4650
3	0.3835	0.2361	0.1247	0.1093	0.0670	0.0546	0.0247	1.6414
4	0.4545	0.2045	0.1212	0.1193	0.0473	0.0341	0.0189	1.5272
5	0.3256	0.1628	0.1550	0.1395	0.1085	0.0543	0.0543	1.7819
Week	$\hat{\boldsymbol{q}}(\texttt{Week})$							$\hat{\mathcal{E}}$
1	0.3955	0.3284	0.1791	0.0522	0.0448	0.0000	0.0000	1.3338
2	0.2925	0.2054	0.1826	0.1535	0.1100	0.0311	0.0249	1.7255
3	0.3419	0.1581	0.1496	0.1026	0.0940	0.0940	0.0598	1.7894
4	0.2654	0.2099	0.1852	0.1636	0.1111	0.0340	0.0309	1.7545
5	0.4508	0.2798	0.1088	0.0622	0.0622	0.0259	0.0104	1.4443

Weekly Estimates

2016 June NCBS vs. 2017 March NCBS

Week	$\hat{\boldsymbol{p}}(\texttt{Week})$							$\hat{\mathcal{E}}$
1	0.4023	0.3456	0.1133	0.0793	0.0595	0.0000	0.0000	1.3492
2	0.4565	0.2798	0.1285	0.0763	0.0469	0.0067	0.0054	1.3793
3	0.4676	0.2444	0.1321	0.0819	0.0647	0.0053	0.0040	1.3989
4	0.5431	0.1802	0.0992	0.0653	0.0522	0.0470	0.0131	1.4022
5	0.4505	0.2344	0.1465	0.1136	0.0403	0.0147	0.0000	1.4190
Week	$\hat{\boldsymbol{q}}(\texttt{Week})$							$\hat{\mathcal{E}}$
1	0.3955	0.3284	0.1791	0.0522	0.0448	0.0000	0.0000	1.3338
2	0.2925	0.2054	0.1826	0.1535	0.1100	0.0311	0.0249	1.7255
3	0.3419	0.1581	0.1496	0.1026	0.0940	0.0940	0.0598	1.7894
4	0.2654	0.2099	0.1852	0.1636	0.1111	0.0340	0.0309	1.7545
5	0.4508	0.2798	0.1088	0.0622	0.0622	0.0259	0.0104	1.4443

Weekly Estimates

June NCBS vs. 2016 September NCBS

Week	$\hat{\boldsymbol{p}}$(Week)							$\hat{\mathcal{E}}$
1	0.4023	0.3456	0.1133	0.0793	0.0595	0.0000	0.0000	1.3492
2	0.4565	0.2798	0.1285	0.0763	0.0469	0.0067	0.0054	1.3793
3	0.4676	0.2444	0.1321	0.0819	0.0647	0.0053	0.0040	1.3989
4	0.5431	0.1802	0.0992	0.0653	0.0522	0.0470	0.0131	1.4022
5	0.4505	0.2344	0.1465	0.1136	0.0403	0.0147	0.0000	1.4190
Week	$\hat{\boldsymbol{q}}$(Week)							$\hat{\mathcal{E}}$
1	0.3857	0.3327	0.2388	0.0245	0.0122	0.0061	0.0000	1.2515
2	0.4224	0.3077	0.1190	0.0581	0.0421	0.0276	0.0232	1.4650
3	0.3835	0.2361	0.1247	0.1093	0.0670	0.0546	0.0247	1.6414
4	0.4545	0.2045	0.1212	0.1193	0.0473	0.0341	0.0189	1.5272
5	0.3256	0.1628	0.1550	0.1395	0.1085	0.0543	0.0543	1.7819