

פרויקט תכנון ניסויים וניתוח שונות

המחלקה להנדסת תעשייה וניהול אוניברסיטת בן גוריון



שם: חאלד סובח

תוכן עניינים

[שאלה 1 3](#_Toc40043069)-6

[תיאור הבעיה והנחות המודל 3](#_Toc40043072)

[בדיקת הנחות המודל 4-5](#_Toc40043078)

[אמידת פרמטרים, ניתוח המודל וניתוחי המשך 5-7](#_Toc40043069)

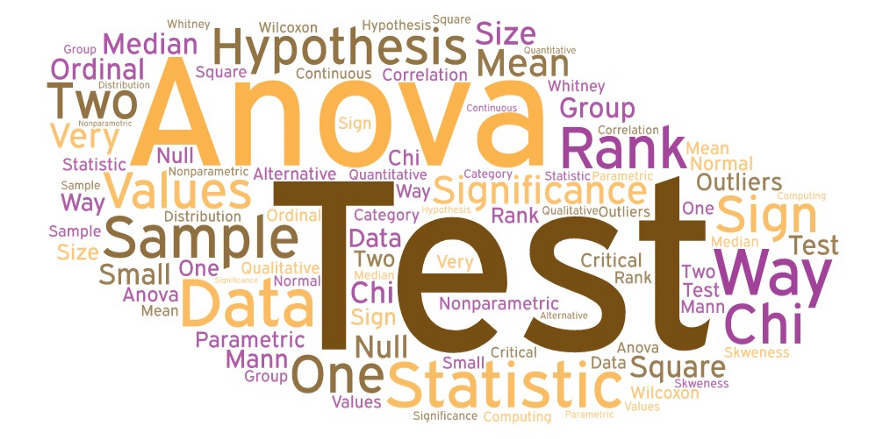
[שאלה 2 8](#_Toc40043072)-11

[סמלוץ נתונים ובדיקת הנחות המודל המסומלץ 8-9](#_Toc40043069)

[ניתוח המודל המסומלץ 9-10](#_Toc40043072)

[[ניתוח סימולציה של 1000 ניתוחים 10-11](#_Toc40043069)](#_Toc40043072)

[[תיעוד קוד](#_Toc40043069)](#_Toc40043072) [12-14](#_Toc40043069)



**תיאור הניסוי:**

סטארבקס משתמשת בקונים סודיים שנראים כלקוחות כדי להיכנס לחנות ולתעד את החוויה שלהם מבחינת שירות לקוחות, ניקיון ואיכות. הקונים הסודיים מקבלים הכשרה סטנדרטית של סטארבקס על מנת להבטיח תמימות ואובייקטיביות בביקורות החנות שלהם.

בגלל מיקומו בערים סידני, בריסביין ומלבורן באוסטרליה, בחרה סטארבקס 6 מהקונים הסודיים שלה באופן אקראי. כל אחד מששת הקונים יוקצה לבקר באותה חנות בכל אחת משלוש הערים.

התוצאות שהתקבלו מהניסוי של החברה:

|  |  |  |  |
| --- | --- | --- | --- |
| Melbourne | Brisbane | Sydney | Shopper\city |
| 88 | 73.6 | 74.3 | Shopper1 |
| 72 | 69.3 | 69.3 | Shopper2 |
| 74.3 | 55 | 54.6 | Shopper3 |
| 87.7 | 64.6 | 65.6 | Shopper4 |
| 80.3 | 65.6 | 80.3 | Shopper5 |
| 65.6 | 65.6 | 66.6 | Shopper6 |

החברה רוצה לדעת האם יש הבדל בערים שנבדקו, ואם כן, החברה רוצה לבדוק יותר את איכות החנות הגרוע ביותר.(***כלומר: שאלת המחקר)***

**יש לנו מודל חד כיווני עם בלוק**. סטארבקס בחרה 3 ערים ספציפיים לבדיקה, ולכן יש לנו גורם ראשי קבוע. ששת הקונים או הבודקים של החברה נבחרו באופן אקראי מבין כל הקונים שעובדים בחברה, ולכן יש לנו בלוק.

**המודל**:

מספר הבלוקים-הקונים

מספר הטיפולים-הערים

- המדידה של הטיפול ה- בבלוק ה-

- הממוצע הכללי של כל האוכלוסייה

- האפקט של הקבוצה ה-

- האפקט של הבלוק ה-

- הטעות האקראית של טיפול ה-בבלוק ה-

הנחות המודל:

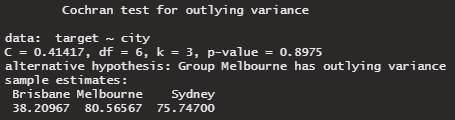
* באים מהתפלגות נורמלית
* בלתי תלויים בין הקבוצות ובתוך הקבוצות

שלב 1:

בדיקת הנחות המודל:

1. הנחת שוויון שונויות:

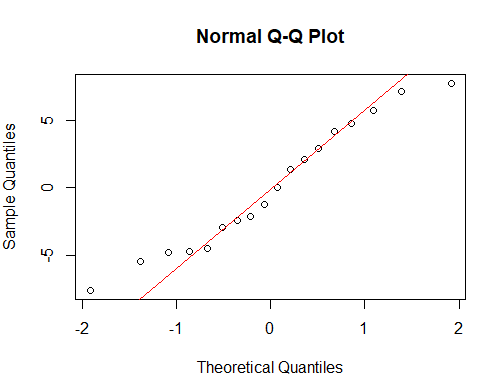
נבדוק את ההשערה בעזרת מבחן קוכרן(אותו מספר של תצפיות בכל קבוצה):



ניתן לראות שה- ולכן לא נדחה את השערת האפס ונגיד שהנחת שוויון השונויות מתקיימת.

\*הערה: בבדיקה ההנחה, לא התייחסתי לבלוק(לא הצלחתי למצוא מידע טוב על זה), אבל מה שאני חושב, זה שאם הבלוק הינו מקור נוסף לשונות, אז זה לא מאוד משפיע אם התעלמנו ממנו אם לא דחינו את ההשערה וההנחה מתקיימת.

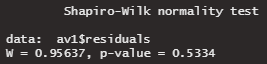
2. הנחת הנורמליות:



ניעזר בתרשים QQ-Plot:

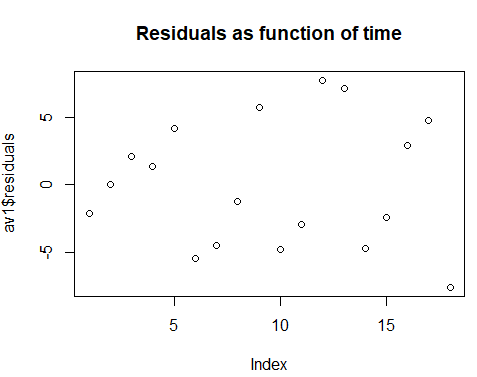
לפי התרשים, נוכל לראות שיש בריחה בזנבות, אבל לרוב נראה שההנחה מתקיימת(התצפיות באמצע). ולכן עדיף להשתמש במבחן סטטיסטי כדי לקבוע האם הנחת הנורמליות מתקיימת או לא.

נשתמש במבחן .W.S כדי לוודא את הטענה שלנו:



ניתן לראות שה- ולכן לא נדחה את השערת האפס ונגיד שהנחת הנורמליות מתקיימת.

3. הנחת אי-תלות:

ניעזר בתרשים שאריות כנגד זמן:

התצפיות מתפזרות סביב האפס במידה סבירה, אין מגמתיות בולטת בתרשים. ולכן אפשר לטעון שהנחת האי-תלות גם מתקיימת.

שלב 2:

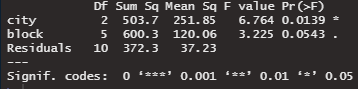
אחרי שבדקנו את הנחות המודל, נבדוק את ההשערה לגבי ההצדקה לשימוש בבלוקים:

השערות לבדיקה:

אמידת פרמטרים:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ממוצע | Melbourne | Brisbane | Sydney | Shopper\city |
| 78.63 | 88 | 73.6 | 74.3 | Shopper1 |
| 70.2 | 72 | 69.3 | 69.3 | Shopper2 |
| 61.3 | 74.3 | 55 | 54.6 | Shopper3 |
| 72.63 | 87.7 | 64.6 | 65.6 | Shopper4 |
| 75.4 | 80.3 | 65.6 | 80.3 | Shopper5 |
| 65.93 | 65.6 | 65.6 | 66.6 | Shopper6 |
|  | 68.45 | 77.98 | 65.62 | ממוצע |

טבלת ANOVA:



בדיקת תכנון המודל: הבלוק לא מובהק(מקור נוסף לשונות), ולכן נעמיס אותו ונגיד שלא היה הצדקה לשימוש בבלוקים.

SSE משתנה ל:

טבלת ANOVA אחרי ההעמסה:

Graphical user interface, text

Description automatically generated

מסכנה: הגורם הראשי(הקבוע) הינו גם מובהק,זאת אומרת שיש הבדל בין איכות החנויות בערים שנבדקים. ולכן נמשיך לניתוחי המשך-נבצע חלוקה לקבוצות הומוגניות.

שלב 3:

ניתוח המשך (Post-Hoc):

יש לנו ניסוי מאוזן, ולכן אפשר להשתמש במבחן HSD:

|  |  |  |  |
| --- | --- | --- | --- |
| Melbourne | Brisbane | Sydney |  |
| 68.45 | 77.98 | 65.62 |  |

S

B

M

אין קבוצה זרה, ולכן:

**מסכנה**: עבור רמת מובהקות 0.05, לא נוכל להמליץ על אף עיר משלושת העירים שבדקנו.

אולי היינו כן יכולים להמליץ על עיר מסוימת שהיא פחות איכותית מבחינת הבדיקה שעשו, אם היינו מורידים את רמת המובהקות, או אם היה לנו יותר בדיקות.

שאלה 2:

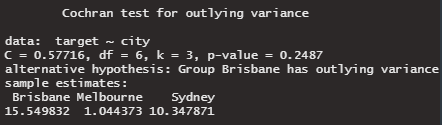
1. נניח שלמנהלי החברה יש חשש שהלקוחות בעיר סידני אינם מרוצים משירות הלקוחות בהשוואה עם העירים בריסביין ומלבורן. ולכן הם סמלצו נתונים על הבעיה לפני הקמת הניסוי האמיתי כדי לדעת אם הם צודקים בטענתם או לא. וקיבלו את התוצאות הבאות:

|  |  |  |  |
| --- | --- | --- | --- |
| Melbourne | Brisbane | Sydney | Shopper\city |
| 71.89 | 73.82 | 67.43 | Shopper1 |
| 73.92 | 78.43 | 66.68 | Shopper2 |
| 70.89 | 67.26 | 63.48 | Shopper3 |
| 71.57 | 75.43 | 66.6 | Shopper4 |
| 71.79 | 74.7 | 64.66 | Shopper5 |
| 72.25 | 77.43 | 72.78 | Shopper6 |

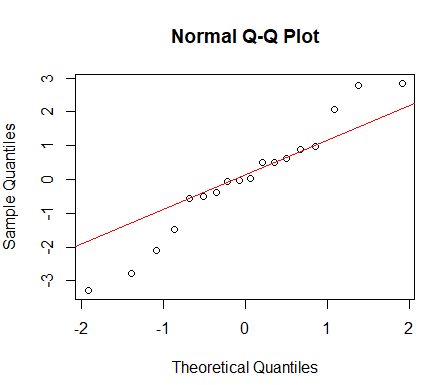
2. בדיקת הנחות המודל:

. הנחת שוויון שונויות:

נבדוק את ההשערה בעזרת מבחן קוכרן(אותו מספר של תצפיות בכל קבוצה):



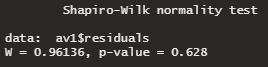
ניתן לראות שה- ולכן kt נדחה את השערת האפס ונגיד שהנחת שוויון השונויות מתקיימת.

2. הנחת הנורמליות:

ניעזר בתרשים QQ-Plot:

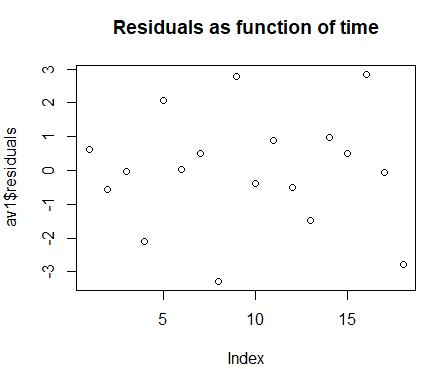
לפי התרשים, נוכל לראות שיש בריחה בזנבות, אבל לרוב נראה שההנחה מתקיימת(התצפיות באמצע). ולכן עדיף להשתמש במבחן סטטיסטי כדי לקבוע האם הנחת הנורמליות מתקיימת או לא.

נשתמש במבחן .W.S כדי לוודא את הטענה שלנו:



ניתן לראות שה- ולכן לא נדחה את השערת האפס ונגיד שהנחת הנורמליות מתקיימת.

3. הנחת אי-תלות:

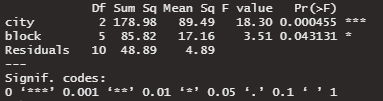
ניעזר בתרשים שאריות כנגד זמן:

התצפיות מתפזרות באופן סביב האפס, אין מגמתיות בולטת בתרשים. ולכן אפשר לטעון שהנחת האי-תלות גם מתקיימת, אבל עדיף לעשות מבחן סטטיסטי לבדיקה נכונות הטענה שלנו.

עכשיו, נבדוק את ההשערה לגבי ההצדקה לשימוש בבלוקים:

השערות לבדיקה:

טבלת ANOVA:



מסכנות:

* הבלוק מובהק: ולכן תכנון הניסוי לפי בלוקים אכן מוצדק ואין צורך להעמסה.
* הגורם הראשי גם מובהק, ולכן נגיד שיש הבדל בין קבוצות הגורם הראשי ונמשיך וניתוחי המשך

ניתוח המשך (Post-Hoc):

יש לנו ניסוי מאוזן, ולכן אפשר להשתמש במבחן HSD:

|  |  |  |  |
| --- | --- | --- | --- |
| Melbourne | Brisbane | Sydney |  |
| 72.06 | 74.52 | 66.94 |  |

S

B

M

יש קבוצה זרה, ולכן:

**מסכנה**: עבור רמת מובהקות 0.05, נוכל להמליץ על העיר סידני. כלומר, איכות השירות בעיר הזאת הינה נמוכה מהערים בריסביין ומלבורן.

3. עכשיו נבצע את הניתוח לגבי טבלת ANOVA, עבור מובהקות הבלוק ומובהקות הגורם הראשי. הסימולציה מבצעת באופן הבא:

1. נסמלץ נתונים מחדש עבור כל ניתוח
2. נבנה טבלת ANOVA ונבדוק: אם הבלוק לא מובהק נבצע העמסה ונבדוק את מובהקות הגורם הראשי. אם אבלוק כן מובהק, נבדוק אם הגורם הראשי מובהק או לא.
3. בכל שלב נרשום אם הבלוק והגורם הראשי היו מובהקים או לא.

חוזרים על התהליך הזה 1000 פעמים, ונקבל את התוצאות הבאות:

|  |  |  |
| --- | --- | --- |
| מובהק | לא מובהק |  |
| 755 | 245 | עיר |
| 33 | 967 | בלוק |

מסכנות הסימולציה:

1. מסכנה על הבלוק: רוב הפעמים (967) הבלוק לא היה מובהק. אפשר להגיד שהתכנון לפי הבלוק הזה אינו מוצדק ואפשר לעבור למודל חד כיווני עם גורם ראשי אחד במידה והנחות המודל מתקיימות.
2. ניתן לראות שב-755 פעמים דחינו את השערת האפס עבור הגורם הראשי, כלומר אכן יש הבדל בין קבוצות הגורם הראשי ונמשיך לניתוחי המשך. ולכן, נמליץ על קיום הניסוי כדי לנתח את הנתונים האמיתיים.

Text

Description automatically generatedתיעוד קוד:

Text

Description automatically generated

Text

Description automatically generated