**השפעת טעויות עיגול גדולות על ביצועים של תרשים בקרה לתוחלת כאשר אופיין האיכות מתפלג נורמלית ושונותו ידועה**

**רקע**

תרשימי הבקרה של Shewhart הם כלי חשוב לניטור שוטף של תהליכים. בכל תהליך ייצור קיימת השתנות טבעית של משתני התהליך הנובעת מסיבות אקראיות. תהליך שבו יש רק סיבות אקראיות להשתנות הוא תהליך הנמצא תחת בקרה סטטיסטית. לעיתים ההשתנות אינה נובעת רק מסיבות אקראיות אלא נובעת גם מסיבות מיוחסות, שמקורן בבעיה במכונה, בעובד, בחומר הגלם או בגורמים משפיעים נוספים. היצרן בוחר את המשתנים אותם ירצה לנטר, משתנים אלו נקראים אופייני איכות (Quality Characteristics). התפקיד העיקרי של תרשימי הבקרה הוא לנטר באופן שוטף את הפרמטרים, כגון התוחלת והשונות, של אופיין האיכות, כשהמטרות העיקריות הן זיהוי מהיר ככל האפשר של סטיות בתהליך, אמידת הפרמטרים של התהליך וקביעת כושר התהליך תוך מזעור אזעקות שווא.

בשימוש בתרשימי בקרה לא נהוג להתייחס לסטיות הנובעות מטעויות מדידה אקראיות ושיטתיות, אולם טעויות מדידה משמעותיות עלולות להשפיע על תרשים הבקרה ובמקרים מסוימים אף לגרום להסקת מסקנות שגויות לגבי בקרת התהליך על ידי הגדלת טעות מסוג ראשון (הסיכוי לאזעקת שווא, החלטה על יציאה מבקרה כאשר אין סטייה בתהליך) או הגדלת טעות מסוג שני (הסיכוי לא לזהות סטייה בתהליך למרות שהיא קיימת).

בספרות המקצועית העוסקת בתרשימי הבקרה הקלאסיים של Shewhart ישנה הנחה שהנתונים הנמדדים הם הערכים האמתיים של המשתנה המנוטר.Bennett (1954) התייחס לנושא במחקרו והציע שכאשר השונות של טעות המדידה קטנה מהשונות של התהליך ניתן להתעלם ממנה כי השפעתה על העוצמה של תרשימי הבקרה קטנה מאוד. Abraham (1977) הציע לחשב את תרשימי הבקרה ללא התייחסות לטעויות המדידה ואז להוסיף ערך קבוע שייצג את טעות המדידה. Kanazuka (1986) הציע לפתור את בעיית היחלשות העוצמה של תרשימי הבקרה במצב של טעויות מדידה על ידי הגדלת המדגם. גם Walden (1990) הציע לפתור את הבעיה בעזרת הגדלת המדגם. פתרון אפשרי נוסף שהציג הוא תהליך של דגימות חוזרות או שילוב של הגדלת המדגם עם דגימות חוזרות. תוצאות אלה ונוספות מוזכרות במאמר הסקירה של Linna and Woodall (2001).

Wheeler (a2011) מתייחס במאמרו לתרשימי בקרה לתכונות כאשר מסווגים מוצר כתקין/פגום באופן שגוי עקב טעויות מדידה. Wheeler מציע במאמרו לשפר את איכות תהליך הייצור ולא להשקיע את המשאבים במערכות מדידה מושלמות. לטענתו, שיפור מערכות המדידה במיון המוצרים כך שיתאימו למפרטים יגדיל את התקורה ואילו שיפור תהליך הייצור יפחית עלויות.

תהליך המדידה נתון להשפעה של גורמים רבים כגון תנאים סביבתיים, הבדלים בין המודדים, טעויות עיגול, שגיאות כיול, שגיאות בקריאת התוצאות ומועדי המדידה. על כן, כאשר מודדים מספר פעמים נמדד כלשהו, סביר להניח שתהיה שונות בין המדידות. ניתן להפריד את הגורמים לשונות במדידות לשני גורמים עיקריים: שונות שמקורה בנמדד עצמו ושונות שנובעת מתהליך המדידה וממכשירי המדידה (Measurement Error). (Gertsbakh, 2003)

ניתן לתאר את הנמדד Y כסכום של שני מרכיבים :

(1)

X- ערכו ה"אמיתי" של הנמדד

ε - הטעות במדידה

ניתן להרחיב את התיאור הכללי של המדידה ולתארו כך:

(2)

- התוחלת של X

- ההפרש בין התוחלת של X לבין ערכו האמתי

- הרעש של הערך האמתי, הטעות האקראית

- טעות המדידה

הנחת היסוד היא אי תלות בין שני מרכיבי הטעות. (Gertsbakh, 2003)

במאמר זה נתייחס לטעות מדידה שנובעת מעיגול. כל מכשיר מדידה מציג מדידות שהן עיגול של התצפיות הנמדדות על פי רמת דיוק המכשיר. טעות העיגול נובעת משתי סיבות עיקריות. הסיבה הראשונה נובעת ממכשיר המדידה עצמו. לעתים, בשל אילוצים טכניים, כלכליים או אחרים, מדידות מבוצעות באמצעות מכשירי מדידה זולים ומהירים עם יחידות מידה גדולות. לדוגמה, אם יחידות המידה של מאזניים הן בקפיצות של יחידות שלמות (למשל קילוגרמים), אז התוצאות שיתקבלו יהיו מספרים שלמים בלבד ולא מספרים עשרוניים, למרות שהערך של המשתנה שנשקל יכול להיות מספר עשרוני (Gertsbakh, 2003). הסיבה השנייה תלויה במערכות הקולטות את תוצאות המדידה. לדוגמה, מכשיר המדידה מחזיר ערך בן 25 ספרות אבל המערכת שקוראת את הנתונים מסוגלת לשמור רק 18 ספרות. Zhidong et al, 2009)).

אחת המשמעויות של עיגול היא שמתקבלות תוצאות מדידה עם ערכים סופיים כתלות ביחידות המידה של המכשור ולא כתלות בערכיו ה"אמתיים" של המשתנה הנמדד. התוצאה היא שמתקבלים נתונים מעוגלים של המשתנה הנמדד. = Y כאשרY מייצג את התצפיות המעוגלות של X. (Gertsbakh, 2003)

עיגול הנתונים משפיע על הניתוחים הסטטיסטיים. המקור להשפעה המשמעותית על ההסקה הסטטיסטית כשהנתונים מעוגלים הוא שבעצם תהליך העיגול קורה תהליך של דיסקרטיזציה לנתונים, כלומר מעבר מטיפול במשתנה מקרי רציף למשתנה מקרי בדיד.

במקרים רבים ניתן להתעלם מהעיגול ללא חשש ולהשתמש באמידה מסורתית של הפרמטרים באמצעות הנתונים המעוגלים, והאומדן נחשב מספיק מדויק לצורך הסקה סטטיסטית. אולם במקרים מסוימים התעלמות מהאופי של תהליך העיגול עלולה להוביל לאי דיוקים משמעותיים באמידה של הפרמטרים ולהביא לכך שהשימוש בכלים הסטטיסטיים התיאורטיים יגרום לשגיאות בהסקה הסטטיסטית. (Gertsbakh, 2003)

מידת העיגול נקבעת לפי היחס בין סטיית התקן של המשתנה הנמדד (σ) לבין יחידת המידה של מכשיר המדידה ((h. היחס מסומן ב- δ *(* ) כאשר יחידת המידה של מכשיר המדידה מוגדרת כהפרש בין שני ערכים עוקבים של מכשיר המדידה ונהוג להניח שהיא ידועה.

ההשפעה של טעות העיגול על ביצועי תרשים הבקרה תלויה בגודלו של δ. ככל שהיחס δ קטן יותר, כך העיגול גס יותר והשפעתו גדולה יותר. (Gertsbakh, 2003)הציג בספרו כללי אצבע לסיווג: כאשר תהליך המדידה ייחשב תהליך רגיל המאפשר שימוש בכלים הסטטיסטיים התיאורטיים המסורתיים. כאשר העיגול ייקרא עיגול גס ותהליך המדידה ייחשב מיוחד, במקרה זה קיים חשש לכך שהשימוש בכלים סטטיסטיים תיאורטיים מסורתיים עלול לגרור להסקה סטטיסטית שגויה ולכן יש לבצע התאמות. כאשר רמת העיגול אינה מקבלת הגדרה מיוחדת וזה נחשב מצב ביניים. (Gertsbakh, 2003)*. את הערך האמיתי של הנמדד,* X *, והתוצאות המעוגלות*Y *ניתן לבטא כך:*

(3)

*- טעות מדידה שמקורה בעיגול גס.*

*מניחים שטעויות המדידה הרגילות ( ) והטעות האקראית (*) *הינן זניחות בהשוואה לטעות עיגול-גס ולכן במשוואה 3 מחליפה את במשוואה 1 ואת במשוואה 2.*התצפיות המתקבלות הן כאשר

באופן תיאורטיY יכול לקבל כל ערך שהוא כפולה של h, אך בפועל Y שנמדד עם טעות מדידה שמקורה בעיגול גס (0.5<δ ) יוכל לקבל לכל היותר 5 ערכים שונים בעלי הסתברות משמעותית (גדולה מ 0.00001). אמנם יתקבלו עוד ערכים, אבל בהסתברויות קטנות מאוד ולכן ניתן להתעלם מהם.

תחת ההנחה שהמשתנה הנמדד לפני עיגול X מתפלג נורמלית , פונקציית ההסתברות של המשתנה המקריY היא:

 (4)

(Gertsbakh, 2003)

– הערך המעוגל השכיח ביותר (בעל ההסתברות הגבוהה ביותר).

Benson el al (2013) חקרו את אמידת השונות כשהתוחלת ידועה והנתונים מתפלגים נורמלית עם טעות מדידה שמקורה בעיגול גס. Benson et al (2014) המשיכו וחקרו את אמידת השונות גם בעזרת רווח בר סמך בנתונים המתפלגים נורמלית עם תוחלת ידועה שנאספו עם טעות מדידה שמקורה בעיגול גס.

מטרתו של מאמר זה היא ללמוד על השפעת טעויות עיגול גס על ביצועי תרשים הבקרה של שוהרט לתוחלת. מבנה המאמר: בפרק 2 יוצגו חישובים תיאורטיים לביצועי תרשימי הבקרה של שוהרט כאשר הנתונים ללא עיגול וכשהנתונים מעוגלים בעיגול גס. בפרק 3 תוצג השוואת ביצועים של תרשימי הבקרה כשהנתונים מעוגלים מול מצב שהנתונים לא מעוגלים. בפרק 4 יוצג ניתוח ברמת המדגם הבודד בעזרת סימולציה. בפרק 5 יוצגו סיכום ומסקנות.

**המודל התיאורטי לביצועי תרשים בקרה כאשר הנתונים ללא עיגול וכשהנתונים מעוגלים בעיגול גס**

על מנת להשוות בין ביצועי תרשים הבקרה כשהנתונים לא מעוגלים לבין ביצועיו כשהנתונים מעוגלים, נערכו חישובים תיאורטיים למדדים הבאים:

אלפא (α) – טעות מסוג ראשון – ההסתברות שממוצע המדגם יהיה מחוץ לגבולות הבקרה כאשר למעשה אין שינוי בתוחלת התהליך.

בטא (β) – טעות מסוג שני – ההסתברות שממוצע המדגם יהיה בתוך גבולות הבקרה כאשר למעשה חל שינוי בתוחלת.

ARL0 – אורך הריצה הממוצע עד לאזעקת שווא

is the expected number of samples until a control chart signals, given that the process is under control

ARL1 – is the expected number of samples until a control chart signals, given the process is out of control (the mean has shifted)

החישובים בוצעו תחת ההנחה שהתצפיות בלתי תלויות והתהליך מתפלג נורמלית עם סטיית תקן ידועה וגודל מדגם קבוע.

בשלב ראשון חושבו גבולות הבקרה עבור נתונים מעוגלים ולא מעוגלים, בהנחה שהתהליך נמצא תחת בקרה.

כשהנתונים לא מעוגלים, המדדים חושבו בעזרת נוסחאות לחישוב הסתברויות של ההתפלגות הנורמלית, כמקובל בתרשימי הבקרה של Shewhart. (Montgomery ,2013)*.*

(5)

ערכי בטא חושבו תחת ההנחה שישנה סטייה של התוחלת בגודל של k סטיות תקן של התהליך. בעזרת גבולות הבקרה שחושבו תחת הנחה שהתהליך מבוקר, חושב הסיכוי לא לזהות את השינוי בתוחלת, כלומר הסיכוי לכך שממוצע המדגם נמצא בין גבולות הבקרה למרות שחל שינוי בתהליך.

פורמלית, יהי

*k- הסטייה של התוחלת מהתוחלת המקורית ביחידות של סטיית תקן, אזי*

(6)

כאשר הנתונים מעוגלים הם **אינם** מתפלגים נורמלית, התפלגותם בדידה כתלות במידת העיגול. לפיכך, חישוב המדדים נעשה בעזרת התפלגות מולטינומית. ראשית חושבה התפלגות הממוצעים () שמהווה בסיס לכל החישובים עבור נתונים מעוגלים. לאחר מכן חושב הממוצע הכללי () וגבולות הבקרה שלפיהם חושבו ערכי ארבעת המדדים: אלפא (α), בטא (β), ARL0 ו ARL1.

התפלגות הממוצעים () חושבה במספר שלבים.

**בשלב הראשון** נבנתה התפלגות הנתונים המעוגלים Y על ידי מציאת גבול עליון וגבול תחתון של ערכי (הנתונים המקוריים ) עבור כל אפשרי (חצי h מכל כיוון), וחושב הסיכוי לקבלת על ידי הפרש הסתברויות מצטברות של הגבולות שנמצאו בעזרת פונקציית ההתפלגות הנורמלית המצטברת.

(7)

התקבלה טבלה המרכזת את 5 ערכי Y בקפיצות של h לכל כיוון מהערך השכיח של Y וההסתברות של כל אחד מהערכים הללו (). יתר הערכים הם בעלי הסתברות קטנה מאוד ואפשר להתעלם מהם.  את הערכים האפשריים של אפשר להציג על ידי ווקטור שמכיל חמישה ערכים (L):

כאשרהוא הערך השכיח (בעל ההסתברות הגבוהה ביותר).

**בשלב השני** נבחר גודל המדגם n, ולכל קומבינציה שלn ערכים אפשריים חושב הממוצע, תחת ההנחה שיש אפשרות לחזרתיות של הערכים במדגם ואין חשיבות לסדר בו דוגמים את הנתונים.

**בשלב השלישי** נבנתה טבלת ההתפלגות של הממוצעים (שכן כמה קומבינציות יכולות לתת אותו ממוצע) – הערכים האפשריים וההסתברות של כל אחד להתקבל. טבלה זו תשמש לבניית תרשים הבקרה.

מספר הקומבינציות (b) נקבע על פי חישוב מספר החלוקות, כלומר חישוב מספר הדרכים לבחור c עצמים מתוך d עצמים עם חזרות ובלי חשיבות לסדר. ישנם c+d-1 מצבים אפשריים, ויש לבחור c מתוכם תוך צמצום המצבים החוזרים על עצמם בסדר שונה. באופן כללי,

באופן ספציפי, נסמן ב - LL את מספר הערכים האפשריים מתוכם יבחרו הערכים לקומבינציה, במקרה של עיגול גס, ישנם לכל היותר חמישה ערכים בעלי הסתברות משמעותית כלומר, LL=5 d=. גודל המדגם הוא מספר העצמים שייבחרו (c=n), נקבל

*כאשר*

LL- אורכו של הווקטור L, LL=5

n- *גודל המדגם*

b- *מספר הקומבינציות*

לדוגמה, כאשר גודל המדגם הוא 7, מספר הקומבינציות יהיה:

נסמן:

i *אינדקס למשתנה מקרי בתוך קומבינציה* 1...n

j *אינדקס* *לקומבינציה בתוך סך הקומבינציות* 1...b

Yi,j –התצפית ה i בקומבינציה ה j

עבור כל קומבינציה (j) יחושב הממוצע (). כמו כן יחושב הסיכוי () על ידי נוסחת ההתפלגות המולטינומית.

כאשר

- אינדקס של הערכים בווקטור L ( 5...1)

- מספר הפעמים שמופיע הערך ה-l של וקטור LL בקומבינציה j.

לדוגמה אם , עבור n=3 h=2.5

הסיכוי לכל ערך בווקטור L

עבור הקומבינציה (7.5,10,10)

בשלב הבא חושב הסיכוי לכל תוצאת אפשרית (תוצאת ממוצע זהה יכולה להיות למספר קומבינציות שונות) על ידי סכימת ערכי של ערכי זהים. התוצאה של השלב הזה היא טבלת ההסתברויות של כל הערכים האפשריים ל . הערכים בטבלה יהיו בקפיצות בגודל . בסה"כ בטבלה יהיו 4n+1 ערכים.

חישוב התוחלת התיאורטית למספרים המעוגלים הוא Benson et al (2013):

חישוב גבולות הבקרה ( ) יהיה על פי הנוסחאות של Shewhart לתרשימי בקרה למשתנים.

A = 3/sqrt(n)

ערכה של אלפא יחושב על ידי סכימה מצטברת של סיכויים לערכי שנמצאים מחוץ לגבולות הבקרה.

לכל רמת עיגול, ערכי בטא יחושבו כתלות בגודל השינוי של התוחלת ביחידות סטיית תקן (k).

אופן חישוב התפלגות הערכים הממוצעים () בהינתן זהה לחישוב שתואר למעלה. ערכה של בטא יחושב על ידי סכימה מצטברת של הסיכויים לערכים שנמצאים בין גבולות הבקרה הקיימים.

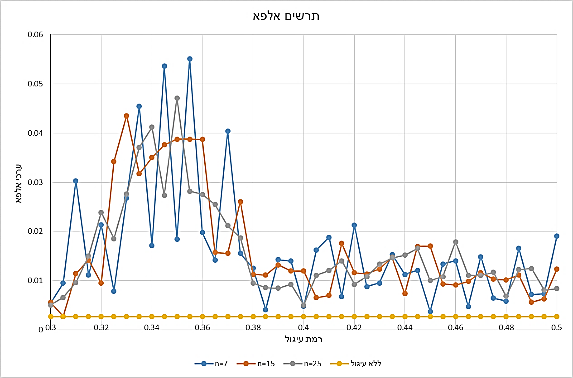
**השוואת הביצועים של תרשימי הבקרה כשהנתונים מעוגלים מול מצב בו הנתונים לא מעוגלים**

בפרק זה יוצגו מספר דוגמאות של החישובים התיאורטיים למדדי ביצוע תרשימי הבקרה כשהנתונים ללא עיגול וכשהנתונים מעוגלים. התוצאות יוצגו עבור ערכים לא מעוגלים מהתפלגות נורמלית עם תוחלת 10 וסטיית תקן 1, ועבור נתונים מעוגלים עם רמות עיגול בטווח בקפיצות של 0.005, סך הכל 41 רמות עיגול. לצורך ההדגמה נבחרו שלושה גדלי מדגם n=7,15 ,25. בדיקת הרגישות של התרשימים נבדקה על ידי 12 ערכי סטיות (k) k=±0.1,0.3,0.5,0.7,1,1.25.

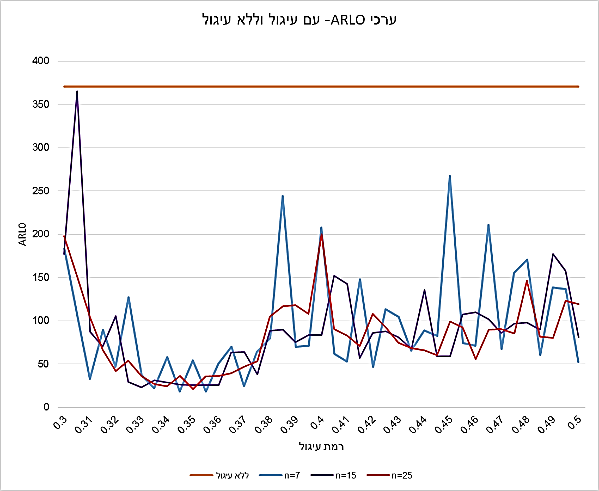
נדגיש שמוצג רק ערך אחד של התוחלת מכיוון שביצועי התרשימים מושפעים מהתוחלת רק דרך הערך של הקו המרכזי שלהם (CL). גם ערכה של סטיית התקן אינו משפיע על ערכי המדדים. בעבור חישובי אלפא, אין משמעות לגודלה של סטיית התקן אלא רק מספר סטיות התקן לפיו נקבעים גבולות הבקרה. בעבור חישובי בטא, יש חשיבות לגודל הסטייה בתוחלת ביחס לגודלה של סטיית התקן ולכן נבדקו ערכי סטייה קטנים, בינוניים וגדולים ביחס לגודלה של סטיית התקן הנבחרת.

**ניתוח תוצאות ערכי אלפא ו ARL0**

בעוד שעבור נתונים לא מעוגלים ערכה של אלפא נקבע ע"י המשתמש וידוע מראש, עבור נתונים מעוגלים ההתפלגות אינה נורמלית וערכה של אלפא אינו קבוע וידוע מראש. בתרשימים הבאים מוצגות השוואות של ערכי אלפא (תרשים ...) וערכי ARL0, אורך הריצה הממוצע עד לאזעקת שווא (תרשים ...), כשהנתונים מעוגלים וללא עיגול.



תרשים X- ערכי אלפא ללא עיגול ועם עיגול כתלות ברמת העיגול ובגודל המדגם



תרשים X- ערכי ARL0 כשהנתונים מעוגלים וללא עיגול כתלות ברמת העיגול ובגודל המדגם

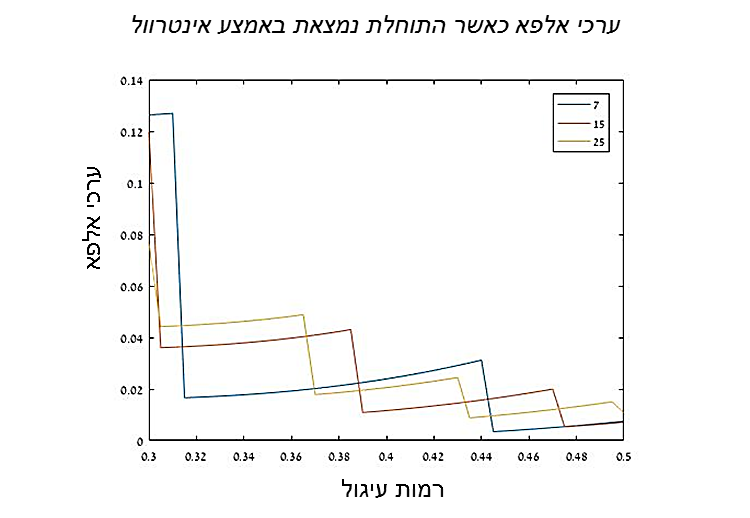
על פי התרשימים לעיל ניתן לראות בבירור שישנה פגיעה בביצועי תרשים הבקרה כשהנתונים מעוגלים. ערכה של אלפא גבוה עבור כל רמת עיגול בטווח הנחקר, עבור כל אחד מגדלי המדגם שנבדקו ובהתאם לכך ערכי ARL0 נמוכים משמעותית מהערך המקובל. המשמעות היא שישנן הרבה יותר אזעקות שווא הפוגעות ברציפות העבודה בפס הייצור וביציבותו.

בתרשימים הללו ניתן לראות שתי תופעות בולטות. ראשית, תופעת השפיצים: עליות וירידות חדות בגרף, תופעה זו מופיעה עבור כל רמות העיגול ובשלושת גדלי המדגם. שנית, ישנו הבדל בערכי אלפא כשרמת העיגול בטווח 0.32-0.38 לבין ערכי אלפא כשרמת העיגול בטווחים 0.3-0.32 ו 0.38-0.5, כאשר הפגיעה בביצועים בולטת יותר כשרמת העיגול היא בטווח 0.32-0.38.

בניסיון להבין את תופעת השפיצים בערכי אלפא בוצע ניסוי חוזר עם **נטרול של המיקום היחסי** של התוחלת בתוך האינטרוול שנוצר כתוצאה מרמת העיגול. נקבע כי התוחלת תהיה בדיוק על הערך המעוגל כלומר התוחלת של X שווה לערך השכיח של Y.

השכיח נקבע בגודל וסביבו נקבעו שאר הערכים האפשריים. L= [,, , , ]

כפי שניתן לראות בתרשים X נמצא שכאשר התוחלת נמצאת בדיוק על הערך השכיח מתקבל גרף המציג התנהגות של שיני מסור בכל אחד משלושת גודלי המדגם.



תרשים X - ערכי אלפא כתלות ברמת העיגול וגדלי המדגם כאשר התוחלת נמצאת באמצע אינטרוול

חישובי אלפא התיאורטיים תלויים ברמת העיגול, בגודל המדגם וגם בגבולות הבקרה של תרשים הבקרה לתוחלת.

לכל רמת עיגול δ יש מספר קבוע של ערכי (ממוצע של קומבינציה בגודל n מבין הערכים האפשריים), עבור גודל סקאלה (h) וגודל מדגם (n) מסוימים. גודל הקפיצה בין ערכי והוא (ההפרש בין שני ערכי סמוכים). לדוגמה עבור n=7 ורמת עיגול δ=0.4 (כלומר h=2.5), גודל הקפיצה בין ערכי הוא .

אם נסמן ב-D את מספר הפעמים שהגודל נמצא בין גבולות הבקרה נקבל:

כאשר ערכו של D שלם אז המשמעות היא שישנה קפיצה, כלומר מתחיל סבב חדש של ערכי ביחס לגבולות הבקרה.לדוגמה, עבור גודל מדגם 7, D יקבל ערך שלם 5 כאשר . בסבב הבא הגודל () נכנס פעם נוספת בכל צד של תרשים הבקרה כך ש D=7. במצב זה הסיכוי להיות בין גבולות הבקרה גדל באופן משמעותי ולכן יש צניחה של אלפא.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **n=7** | **n=15** | **n=25** |
| **D** | **δ** | **δ** | **δ** |
| **3** | 0.1890 | 0.1291 | 0.1000 |
| **5** | 0.3150 | 0.2152 | 0.1667 |
| **7** | 0.4410 | 0.3012 | 0.2333 |
| **9** | 0.5669 | 0.3873 | 0.3000 |
| **11** | 0.6929 | 0.4734 | 0.3667 |
| **13** | 0.8189 | 0.5594 | 0.4333 |
| **15** | 0.9449 | 0.6455 | 0.5000 |
| **17** | 1.0709 | 0.7316 | 0.5667 |
| **19** | 1.1969 | 0.8176 | 0.6333 |
| **21** | 1.3229 | 0.9037 | 0.7000 |

טבלה X- רמות העיגול לפי גדלי מדגם בהם קיימים סבבים מלאים של ערכי כשהתוחלת באמצע האינטרוול

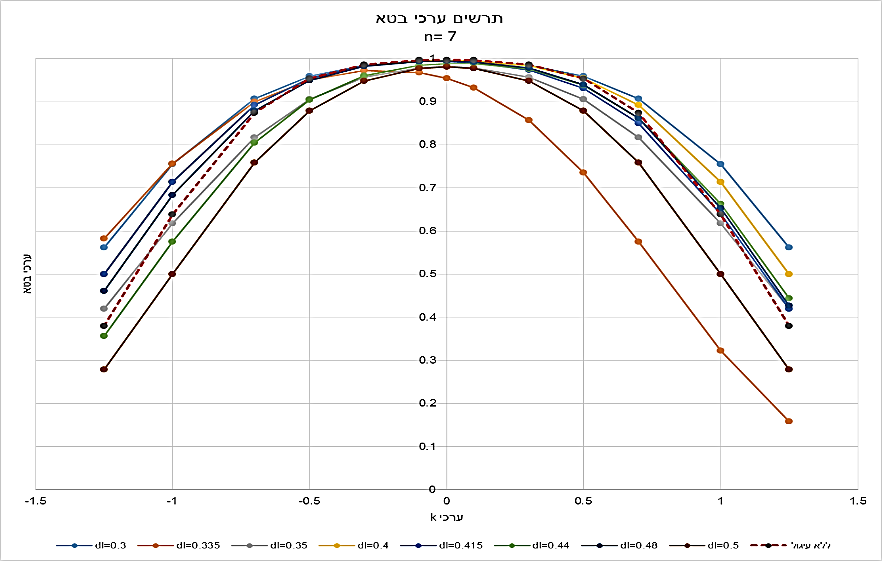
*בטבלה לעיל ניתן לראות את ערכי רמות העיגול בהם מתקיים מצב של סבבים מלאים של ערכי*  בטווח שבין גבול הבקרה והציר המרכזי בעבור גדלי המדגם השונים. ניתן *ללמוד כי בטווח רמות העיגול הנחק*ר , מספר הפעמים בהם מצב זה מתקיים גדל יחד עם גודל המדגם (כשn=7 יש 2 מצבים של ירידה לעומת n=15 בו ישנם 3 מצבים של ירידות וכאשר n=25 בו ישנם 4 מצבים של ירידות). **בתרשים** X שתי הקפיצות בערכי אלפא בתחום הנחקר עבור גודל מדגם 7, הן ברמות עיגול 0.3150 ו-0.4410 כפי שמצוין בטבלה X.

**ניתוח תוצאות ערכי בטא ו ARL1**

בתרשימים הבאים יוצגו ערכי בטא עבור נתונים לא מעוגלים ונתונים המעוגלים ברמות העיגול שונות, בהתאם לגדלי המדגם ולגודל הסטיות. נערכו שתיים עשרה בדיקות רגישות עבור ששה גדלי סטיות k=±0.1,0.3,0.5,0.7,1,1.25.

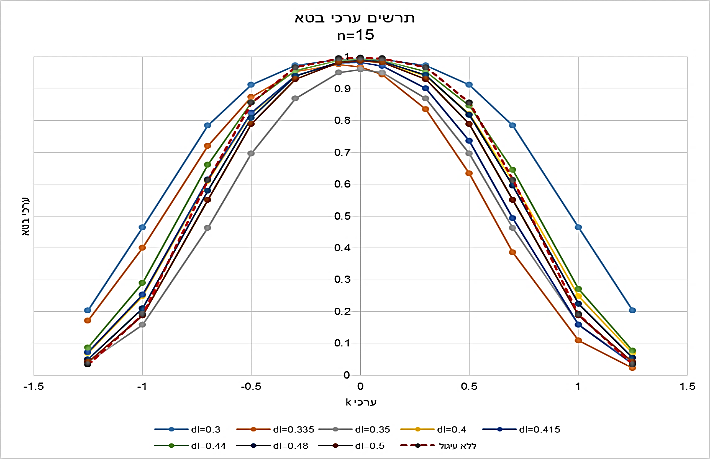
ערכי בטא עבור נתונים מעוגלים ולא מעוגלים

גודל מדגם קטן n=7



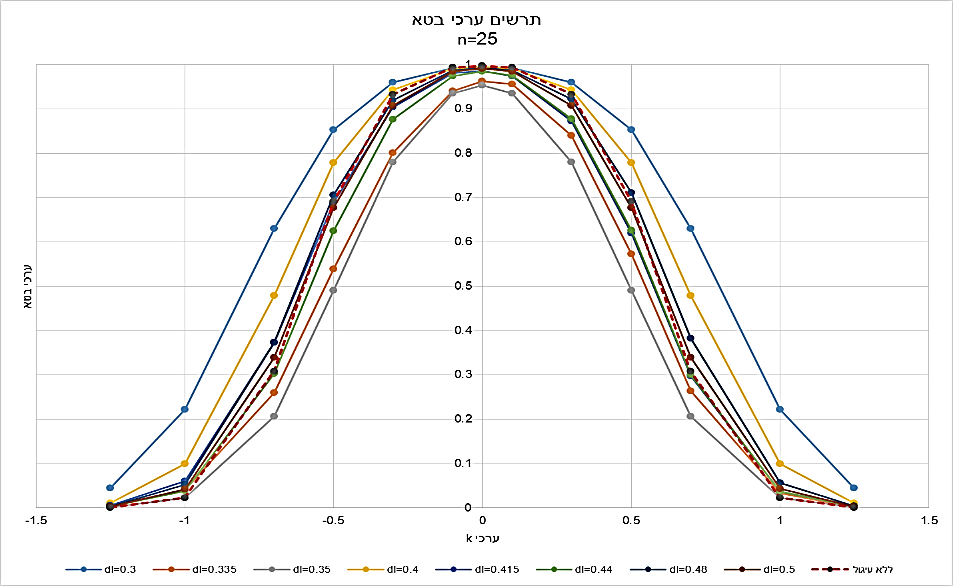
תרשים X- ערכי בטא כשהנתונים לא מעוגלים ומעוגלים כתלות בגודל הסטייה וברמת העיגול n=7

גודל מדגם בינוני n=15



תרשים X- ערכי בטא כשהנתונים לא מעוגלים ומעוגלים כתלות בגודל הסטייה וברמת העיגול n=15

גודל מדגם גדול n=25



תרשים X- ערכי בטא כשהנתונים לא מעוגלים ומעוגלים כתלות בגודל הסטייה וברמת העיגול n=25

בתרשימים אלו מוצגים ערכי בטא עבור 8 רמות עיגול שונות יחד עם ערכי בטא עבור נתונים ללא עיגול (מסומן באדום), בעבור 3 גדלי מדגם. בשלושת התרשימים ניתן לראות כי עבור חלק מהמקרים ערכי בטא כשהנתונים מעוגלים קטנים יותר מערכי בטא כשהנתונים ללא עיגול (ערכים הנמצאים מתחת לתרשים של בטא ללא עיגול – קו אדום מקווקו), משמע ביצועי תרשים הבקרה השתפרו כשהנתונים מעוגלים. לעומתם עבור חלק אחר מהמקרים ערכי בטא כשהנתונים מעוגלים גדולים יותר מערכי בטא כשהנתונים ללא עיגול (ערכים הנמצאים מעל התרשים של בטא ללא עיגול – קו אדום מקווקו), משמע ביצועי התרשים ירדו כשהנתונים מעוגלים. חשוב לציין, ויפורט בהמשך, שהרווח באלפא הוא משמעותית גדול יותר מההפסד בביתא.

כאשר הנתונים אינם מעוגלים קיימת סימטריה מלאה בערכי בטא בעבור סטיות מהתוחלת בגודל זהה בסימנים הפוכים. כשהנתונים מעוגלים בעבור חלק מרמות העיגול הסימטריה נשמרת ואילו עבור חלק מרמות העיגול הסימטריה נפגמת.

**השינוי ב ARL0 לעומת השינוי ב ARL1**

בטבלאות הבאות יוצג השינוי ב ARL0 לעומת השינוי ב ARL1 לפי גודל המדגם עבור 12 בדיקות הרגישות שנעשו. השינוי חושב על ידי הפרש בין ערכי ARL ללא עיגול לבין ערכי ARL כשהנתונים מעוגלים. בעבור ARL0 הפרש חיובי הוא ירידה בביצועי התרשים, מכיוון שישנם פחות מדגמים בין אזעקות שווא כשהנתונים מעוגלים והפרש שלילי הוא שיפור בביצועי התרשים. לעומתו, בעבור ARL1, המשמעות הפוכה, הפרש חיובי הוא שיפור בביצועי תרשים הבקרה (נדרשים פחות מדגמים לזיהוי יציאה אמיתית מבקרה סטטיסטית מאשר כשהנתונים אינם מעוגלים) ואלו הפרש שלילי משמעותו ירידה בביצועי התרשימים (נדרשים יותר מדגמים עד לזיהוי יציאה אמיתית מבקרה כשהנתונים מעוגלים).

|  |  |  |
| --- | --- | --- |
|  | ARL0 | ARL1 |
| הפרש חיובי | ירידה בביצועי תרשימי הבקרה | שיפור בביצועי תרשימי הבקרה |
| הפרש שלילי | שיפור בביצועי תרשימי הבקרה | ירידה בביצועי תרשימי הבקרה |

טבלה X- טבלת משמעות השינוי ב ARL0 וב ARL1

בטבלה הבאה מוצגות התוצאות עבור רמת עיגול 0.42 ושלושה גודלי מדגם.



טבלה X- השינוי ב ARL0 והשינוי בARL1 רמת עיגול 0.42 וגדלי מדגם 7,15,25

ניתן ללמוד שהשינוי ב ARL0 הוא בסדר גודל משמעותי ובכיוון אחד בלבד: יהיו יותר אזעקות שווא בתהליך הניטור כשהנתונים מעוגלים. לעומת זאת השינוי ב ARL1, הינו קטן יותר ולשני הכיוונים. כאשר הסטייה מהתוחלת קטנה מאוד (k=±0.1,0.3), יש שיפור בביצועי התרשים, כלומר הסטייה תתגלה מהר יותר, אך גם במקרים הללו, השינוי קטן ביחס לשינוי ב ARL0.

**ניתוח ברמת המדגם הבודד בעזרת סימולציה**

בפרק הקודם נערכו השוואות באופן כללי. בפרק הנוכחי נשווה ברמת המדגם הבודד. מטרת הסימולציה היא להשוות את אחוז המדגמים בהם תוצאות תהליך הניטור זהות כשהמספרים ללא עיגול וכשהמספרים מעוגלים אל מול אחוז המדגמים בהם תוצאות תהליך הניטור הפוכות. כלומר, מקרים בהם ללא עיגול ממוצע המדגם נמצא בין גבולות הבקרה ואילו כשהנתונים מעוגלים הממוצע נמצא מחוץ לגבולות הבקרה ולהיפך.

ישנם שני מצבים אפשריים, i- ממוצע המדגם נמצא בין גבולות הבקרה, o- ממוצע המדגם נמצא מחוץ לגבולות הבקרה. בטבלה X האות הראשונה (משמאל) מייצגת את המקרה כשהנתונים ללא עיגול והאות השניה מיייצגת את המקרה כשהנתונים מעוגלים. לדוגמה הסימון io מייצג את המצב בו ממוצע המדגם נמצא בין גבולות הבקרה כשהנתונים לא מעוגלים, אך כשהנתונים מעוגלים הוא נמצא מחוץ לגבולות.

|  |  |  |
| --- | --- | --- |
| ללא עיגול- בין הגבולות | ללא עיגול- מחוץ לגבולות |  |
| io | oo | עם עיגול- מחוץ לגבולות |
| ii | oi | עם עיגול – בין הגבולות |

טבלה X- טבלת מצבים בסימולציה לאמידת ביצועי תרשימי הבקרה ברמת המדגם הבודד

מצבים ii,oo מייצגים את המקרים בהם אין השפעה של טעות העיגול (כלומר – ההחלטה זהה).

מצב io מייצג את המקרים בהם כשהנתונים ללא עיגול, ממוצע המדגם נמצא בין גבולות הבקרה (i) ואילו כשהנתונים עוגלו הוא נמצא מחוץ לגבולות הבקרה ((o. תחת H0 מצב זה מעיד על הגדלת אלפא (ירידה בביצועי תרשים הבקרה) כשהנתונים מעוגלים. תחת H1, מצב זה משפר את ביצועי תרשים הבקרה כשהנתונים מעוגלים כלומר מקטין את בטא.

מצב oi מייצג את המקרים בהם ממוצע המדגם נמצא מחוץ לגבולות הבקרה כשהנתונים ללא עיגול ואילו כשהנתונים עוגלו הוא נמצא בין גבולות הבקרה. תחת H0 מצב זה מעיד על הקטנת אלפא (שיפור בביצועי תרשים הבקרה) כשהנתונים מעוגלים. תחת H1, מצב זה פוגע בביצועי תרשים הבקרה כשהנתונים מעוגלים כלומר מגדיל את בטא.

ערכה של אלפא כשהנתונים **לא מעוגלים** הוא הסכום בין שני המצבים בהם ממוצע המדגם נמצא מחוץ לגבולות הבקרה כשהנתונים לא מעוגלים למרות שהתהליך תחת בקרה סטטיסטית (סכום הערכים בעמודה השמאלית בטבלה).

α=oo+oi

ערכה של אלפא כשהנתונים **מעוגלים** הוא הסכום בין שני המצבים בהם ממוצע המדגם נמצא מחוץ לגבולות הבקרה כשהנתונים מעוגלים למרות שהתהליך תחת בקרה סטטיסטית (סכום הערכים בשורה הראשונה בטבלה).

αr=oo+io

ערכה של בטא כשהנתונים **לא מעוגלים** הוא הסכום בין שני המצבים בהם ממוצע המדגם נמצא בין גבולות הבקרה כשהנתונים לא מעוגלים למרות שהתהליך אינו תחת בקרה סטטיסטית (סכום הערכים בעמודה הימנית בטבלה).

β=ii+io

ערכה של בטא כשהנתונים **מעוגלים** הוא הסכום בין שני המצבים בהם ממוצע המדגם נמצא בין גבולות הבקרה כשהנתונים מעוגלים למרות שהתהליך אינו תחת בקרה סטטיסטית (סכום הערכים בשורה השניה בטבלה).

βr=ii+oi

|  |  |  |
| --- | --- | --- |
|  | io גדול | oi גדול |
| H0 | אלפא גדלה  ירידה בביצועי תרשים הבקרה | אלפא קטנה  שיפור בביצועי תרשי הבקרה |
| H1 | בטא קטנה  שיפור בביצועי תרשי הבקרה | בטא גדלה  ירידה בביצועי תרשי הבקרה |

טבלה X- טבלת משמעות המצבים ioו oi תחת H0,H1

שלבי הסימולציה:

תחת השערת האפס H0

1. הגרלת m סטים (m=100K) בגדלים שונים של מספרים מהתפלגות נורמלית בעלת תוחלת וסטיית תקן ידועות.

נבדקו שבעה גדלי מדגם לצורך בדיקת השפעת גודל המדגם על ביצועי התרשימים.

n = [7,10,15,20,25,30,40]

1. חישוב גבולות בקרה (UCL,LCL) בתרשים .

עבור כל אחד משבעת מסדי הנתונים (גדלי n שונים) נבנו גבולות הבקרה שחושבו לפי 3 סטיות תקן.

1. בדיקה עבור כל מדגם j מתוך m המדגמים האם נמצא בין גבולות הבקרה או לא וריכוז התוצאות הבינריות בטבלה בגודל m שורות (כל שורה מייצגת את מספר המדגם). אם אז יישמר הערך 1, אחרת ישמר הערך 0.
2. עיגול סט הנתונים מסעיף א לפי רמות δ שונות
3. אמידת התוחלת למספרים המעוגלים על פי האמד הנאיבי- מיצוע שלn המספרים המעוגלים.

* חישוב גבולות בקרה חדשים (LCLr,UCLr ) לתוחלת (יחושבו כמרחק 3 סטיות תקן מהאמד לתוחלת).

A = 3/sqrt(n)

עבור כל מדגם j מתוך m המדגמים יבדק האם נמצא בין גבולות הבקרה או לא, הוספת התוצאות הבינריות לטבלה מסעיף ג (הוספת עמודות לפי כמות רמות העיגול). אם אז יישמר הערך 1, אחרת ישמר הערך 0.

1. עבור כל סט נתונים בוצעה השוואה של התוצאות ללא עיגול מול התוצאות עם עיגול. נבחנו שיעור ההסכמות בין שני המצבים (מצבים ii,oo), ללא עיגול ועם עיגול, ושיעור המקרים בהם אין ההסכמות (מצבים io,oi). הסימולציה בוצעה כאשר התהליך נמצא תחת בקרה סטטיסטית (תחת H0) וגם כאשר התהליך יצא מבקרה סטטיסטית (תחת H1).

תחת ההשערה האלטרנטיבית H1

עבור כל גודל מדגם הוגרלו שמונה סטים חדשים של m מדגמים כל אחד מהתפלגות נורמלית עם תוחלת שהוסטה בסדרי גודל של סטיית התקן.

נבדקו תשעה ערכיk , שהוצבו כחיוביים ושליליים.

k = [0,0.25,0.3,0.4,0.5,0.75,0.9,1,1.25,1.5]

חזרה על שלבים ב-ו עבור כל סט נתונים.

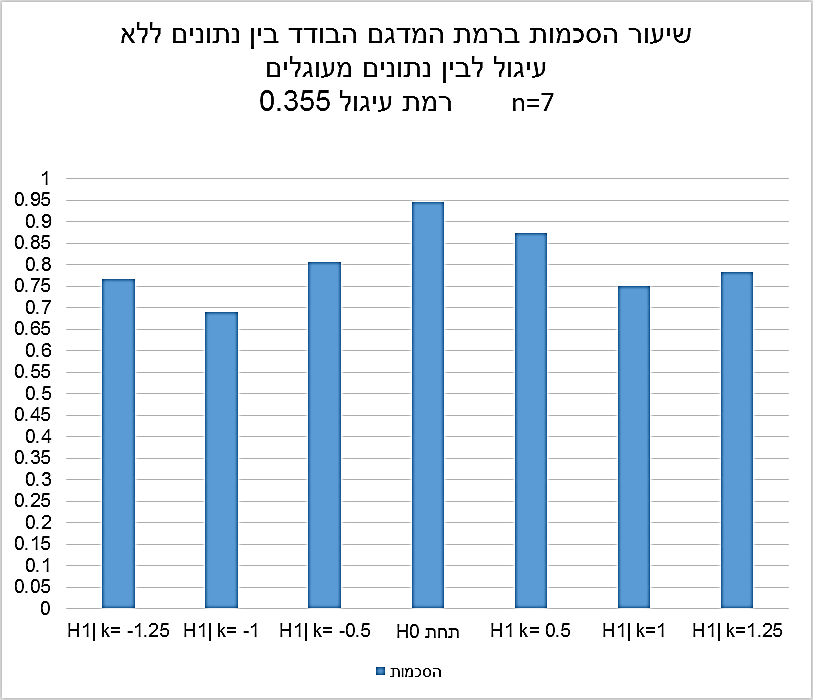
הניתוח יוצג בעבור רמת עיגול δ= 0.355 וגודל מדגם n=7

רמת עיגול 0.355 גודל מדגם 7



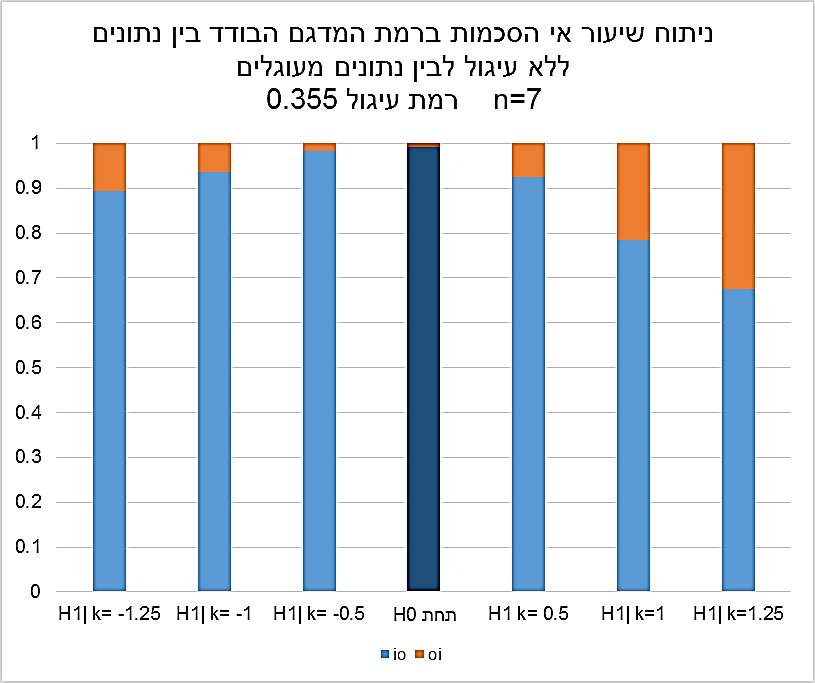
טבלה X- התפלגות המצבים תחתH0,H1 וחישוב ערכי אלפא ובטא, רמת עיגול 0.355 וגודל מדגם 7

מטבלת התוצאות ניתן ללמוד שערכה של אלפא כשהנתונים מעוגלים גדל בצורה משמעותית ביחס לערכה ללא עיגול (פי 22) ואלו ערכה של בטא כשהנתונים מעוגלים יורד ירידה פחות משמעותית ביחס לערכה ללא עיגול (ירידה של 20-50 אחוז).



תרשים X- שיעור הסכמות בין נתונים ללא עיגול לנתונים מעוגלים, רמת עיגול 0.355 וגודל מדגם 7

עבור מדגם קטן ורמת עיגול 0.355 לא ניתן לקבוע את המגמה של שיעור ההסכמות ביחס לרמת הסטייה מהתוחלת (בערך מוחלט). תחת H0 שיעור ההסכמות גבוה (95%) ואילו תחתH1 יורד ואחכ עולה. שיעור ההסכמות נע בין 70-90%.



תרשים X- שיעור אי הסכמות בין נתונים ללא עיגול לנתונים מעוגלים, רמת עיגול 0.355 וגודל מדגם 7

ניתן ללמוד מהתרשים כי עבור מדגם קטן ורמת עיגול 0.355 תחת H0 ישנה ירידה בביצועי התרשימים. בכמעט 100% מאי ההסכמות עבור מיקומו של המדגם ביחס לגבולות נמצא כי המדגמים נמצאים מחוץ לגבולות הבקרה כאשר הנתונים מעוגלים. כלומר ערכה של האלפא גדל כשהנתונים מעוגלים (מצב io).

כאשר יש שינוי בתוחלת (תחת H1), נמצא כי ברוב המקרים כשאין הסכמות בין התוצאות ללא ועם עיגול, המדגם נמצא מחוץ לגבולות הבקרה כשהנתונים מעוגלים כלומר ערכה של בטא קטן (מצב io) וביצועי התרשימים משתפרים.

נראה כי כשהסטייה מהתוחלת חיובית ישנם פחות מצבים בהם המדגמים נמצאו מחוץ לגבולות הבקרה כשהנתונים מעוגלים, אולם גם כאן על פי רוב ניתן לראות כי ערכה של בטא קטן.

ניתן לקבוע כי עבור גודל מדגם קטן ורמת עיגול 0.355 ישנה פגיעה בביצועי תרשים הבקרה עבור אלפא, אולם עבור בטא יש שיפור.

**סיכום ממצאי סימולציה ברמת המדגם הבודד**

* אחוז ההסכמות על מיקום המדגם ביחס לגבולות הבקרה כשהנתונים ללא עיגול וכשאותם הנתונים מעוגלים הינו אחוז גבוה מאוד, במיוחד תחת H0, אולם עדיין ישנו אחוז של אי הסכמות שמשפיע על ביצועי תרשימי הבקרה.
* בעבור המקרים שנבדקו נמצא כי כשגודל המדגם קטן שיעור ההסכמות קטן ככל שהסטייה מהתוחלת גדלה או שישנה ירידה ואז עליה (לרוב עלייה קטנה יחסית) בשיעור ההסכמות כשהסטייה גדלה. בעבור גודל מדגם גדול נמצא כי שיעור ההסכמות הולך וגדל ככל שגודל הסטייה מהתוחלת גדל.
* בעבור המקרים בהם לא היו הסכמות על מיקום המדגם ביחס לגבולות הבקרה, ללא שינוי בתוחלת (תחת H0), בכל המקרים שנבדקו רוב מוחלט של המדגמים נמצא מחוץ לגבולות הבקרה כאשר הנתונים מעוגלים, כלומר ערכה של אלפא גדל באופן משמעותי וחלה ירידה בביצועי תרשים הבקרה.
* בעבור המקרים בהם לא היו הסכמות על מיקום המדגם ביחס לגבולות הבקרה, כאשר יש שינוי בתוחלת (תחת H1) לא ניתן לקבוע באופן גורף על מצבה של בטא כאשר הנתונים מעוגלים. נראו מקרים בהם רוב המדגמים היו בין גבולות הבקרה (ערכה של בטא גדל, ירידה בביצועי תרשים הבקרה), נראו מקרים בהם רוב המדגמים היו מחוץ לגבולות הבקרה (ערכה של בטא קטן, שיפור בביצועי תרשים הבקרה) ואף נראו מצבים של כמעט שיוויון. בכל המקרים השינוי לטובה או לרעה בערכה של בטא הוא שינוי קטן יחסית לשינוי באלפא.
* נמצא כי בעבור סטיות קטנות לרוב יש שיפור בביצועי התרשים, כלומר הסטייה מתגלה יותר מהר כשהנתונים מעוגלים ואילו בעבור סטיות בינוניות וגדולות לרוב יש ירידה בביצועי התרשימים. מבקר האיכות מצפה לקבל התרעה על יציאה מבקרה סטטסטית כשיש סטייה מהתהליך, בעבור סטיות גדולות הוא מצפה לקבל את ההתרעה מהר יותר מאשר עבור סטיות קטנות. כשהנתונים מעוגלים קורה תהליך הפוך, זיהוי היציאה מבקרה סטטיסטית מתארך עבור סטיות גדולות ומתקצר עבור סטיות קטנות.

ממצאי הסימולציה ברמת המדגם מתאימים לממצאי ניתוח תרשימי הבקרה ומדדי אלפא ובטא.

**סיכום ומסקנות כולל מחקר המשך (שיטה חדשה)**

בעבור מדד אלפא ומדד ARL0 ישנה פגיעה משמעותית בביצועי תרשימי הבקרה כאשר הנתונים מעוגלים וגבולות הבקרה מחושבים על פי הכלים של Shewhart. המשמעות הישירה היא שתהיינה אזעקות שווא תכופות יותר שיפריעו לשגרת העבודה של פס הייצור. נזק נוסף שעלול להיווצר עקב ריבוי אזעקות שווא הוא התעלמות של מבקרי האיכות בפס הייצור מההתראות המתקבלות בתהליך הבקרה.

עבור מדד בטא ומדד ARL1 אין השפעה חד משמעית. עבור חלק מרמות העיגול ישנה ירידה בביצועי תרשים הבקרה, כלומר יידרשו יותר מדגמים כדי לזהות יציאה מבקרה סטטיסטית מאשר במצב בו הנתונים ללא עיגול, ובעבור חלק מרמות העיגול יש דווקא שיפור בביצועי התרשימים, כלומר יידרשו פחות מדגמים כדי לזהות את החריגה מאשר במצב בו הנתונים ללא עיגול. מהניתוחים שבוצעו נמצא כי בעבור מצבים בהם רמת הסטייה מהתוחלת קטנה ישנו שיפור, כלומר החריגה תתגלה יותר מהר ודווקא כשהסטייה גדולה או בינונית ומצופה לגלות יחסית מהר את היציאה מבקרה אז ביצועי התרשימים יורדים ונדרשים יותר מדגמים לזיהוי החריגה.

הממצא העיקרי והחשוב הוא שכאשר הנתונים מעוגלים וגבולות הבקרה מחושבים על פי התיאוריה הקלאסית של Shewhart ביצועי התרשימים תחת H0 יורדים בסדר גודל משמעותי עבור אלפא ו ARL0 ביחס לערכם כשהנתונים ללא עיגול. תחתH1 ביצועי התרשימים יורדים או משתפרים בהתאם לגודל הסטייה מהתוחלת, אך בכל המקרים השינוי בערכי בטא ו ARL1 הוא שינוי קטן יחסית לשינוי תחת H0.

ממצא חשוב נוסף הוא היעדר כיוון חד משמעי בשינוי בערכי בטא כאשר הנתונים מעוגלים. בניתוח ברמת המדגם הבודד נראה שהתוצאה שהתקבלה אינה חד משמעית עבור כל המדגמים. במקרה בו בטא גדלה, בניתוח ברמת המדגם הבודד נמצאו אחוזים לא מבוטלים בהם ההתנהגות של חלק מהמדגמים הייתה הפוכה. ממצא זה מעיד על כך שהמימוש בתרשימי הבקרה הסטנדרטיים אינו מתאים עבור נתונים המעוגלים באופן גס.

**ביבליוגרפיה**

1. Abraham, B. (1977). “Control Charts and Measurement Error”. *Annual Technical Conference of the American Society for Quality Control*, 31, pp. 370–374.
2. Bennett, C. A. (1954). “Effect of Measurement Error on Chemical Process Control”. *Industrial Quality Control*, 10, pp. 17–20.
3. Benson, D., Dvir-Harcabi, E., Regev, I. and Schechtman E. (2013). “Estimation of a normal process variance from measurements with large round-off errors”. *IET Science, Measurement and Technology*, 7 (3), pp. 180-189.
4. Benson, D, Schechtman E. (2015). “Using Measurements with Large Round-Off Errors Interval Estimation of Normal Process Variance.” . *IET Science, Measurement and Technology*, pp. 1-7.
5. Gertsbakh, I. (2003). *Measurement Theory for Engineers* (Springer - Verlag, Berlin Heidelberg).
6. Kanazuka, T. (1986). “The Effects of Measurement Error on the Power of X̄ - R Charts”. *Journal of Quality Technology*, 18, pp. 91–95.
7. Linna, K., Woodall, W.(2001), "Effect of measurement error on Shewhart control charts", *Journal of Quality Technology*, pg. 213-222
8. Montgomery, D.C. (2013). *Statistical Quality Control: A Modern Introduction, 7th ed, International Student Version* (Wiley, New York).
9. Vardeman, S.B. (2005). "Sheppard's correction for variances and the quantization noise model", *IEEE Transactions on Instrumentation and Measurement*, Vol. 54, pp 2117 - 2119.
10. Walden, C. T. (1990). “An Analysis of Variables Control Charts in the Presence of Measurement Error”. Unpublished Master's Thesis, Department of Industrial Engineering, Mississippi State University, UMI No. 1343291.
11. Wheeler, D. (2011a), 100% Inspection and Measurement Error, https://www.qualitydigest.com/inside/quality-insider-column/100-inspection-and-measurement-error.html
12. Zhidong B., Shurong Z., Baoxue Z., Guorong H.,(2009). "Statistical analysis for rounded data", *Journal of Statistical Planning and Inference*,139 2526-2542.