משפט הגבול המרכזי

אם מתוך אוכלוסייה בעלת תוכלת מיו ושונות סיגמא בריבוע היינו מוציאים את כל המדגמים האפשריים באותו גודל N (גודל המדגם) ובכל מדגם מחשבים את ממוצע המדגם אזי סדרת ממוצעי כל המדגמים עבור N מספיק גדול, שואפת להתפלגות נורמלית עם תוכלת מיו ושונות סיגמא בריבוע לחלק לN – סטיית התקן נקראת טעות תקן סיגמא לחלק לשורש N.

תאכלס – אם יש מדגם גדול לאללה אז כן זה מתפלג נורמלית – בקורס שלנו מעל 30 ממוצעים נחשב גדול מספיק ואז מבחינתנו זה מתפלג נורמלית.

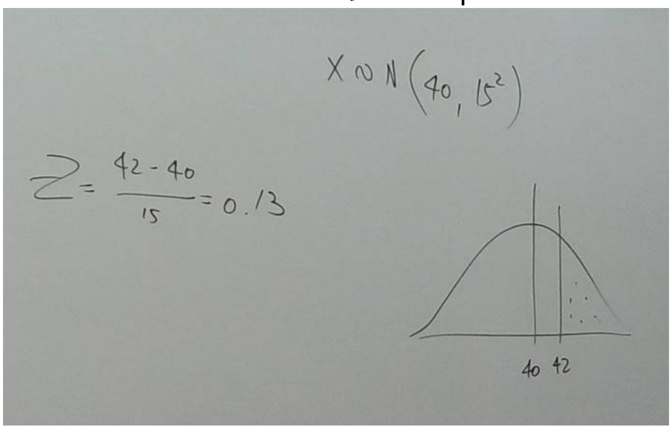
1. אם ידוע שהערכים מתפלגים נורמלית , אזי הממוצעים יתפלגו נורמלית.
2. אם לא ידוע איך הערכים מתפלגים, או לחילופין ידוע שהערכים מתפלגים לא נורמלית, אזי אם N גדול או שווה ל30 אז מבחינתנו הממוצעים יתפלגו נורמלית (משפט הגבול המרכזי)
3. אם לא ידוע איך הערכים מתפלגים וN קטן מ30 – צריך לבדוק איך הערכים מתפלגים (מבחן סטטיסטי שלומדים בסמסטר ב')

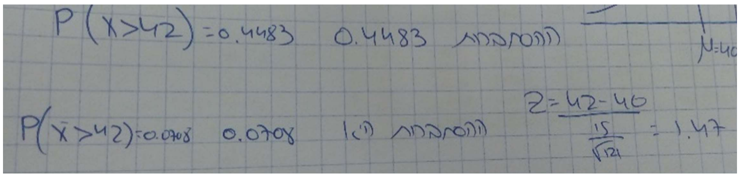
תאכלס אם אני מבין שזה לא נורמלי לזרום כאילו זה נורמלי ורק לציין שצריך לבדוק אם זה אכן נורמלי במבחן שלא למדנו איך לעשות אותו עדיין.

דוגמה:

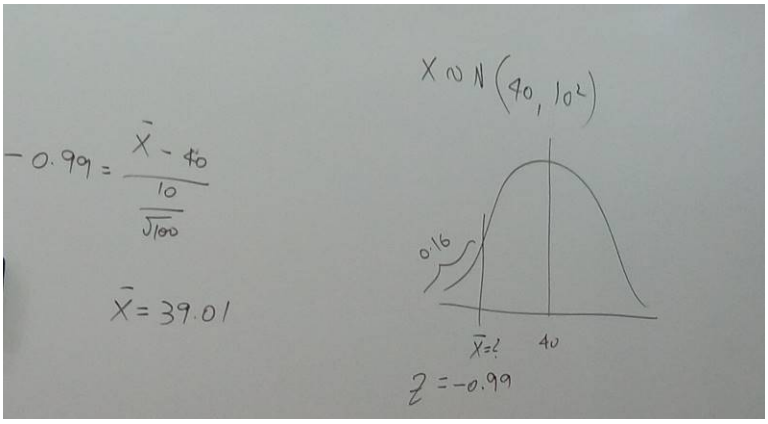
אוכלוסייה שמתפלגת נורמלית, נדגם מדגם מקרי של 121 מתוך האוכלוסייה,

א. מה ההסתברות לקבל ערך גדול מ42
ב. מה ההסתברות לקבל ממוצע גדול מ42?

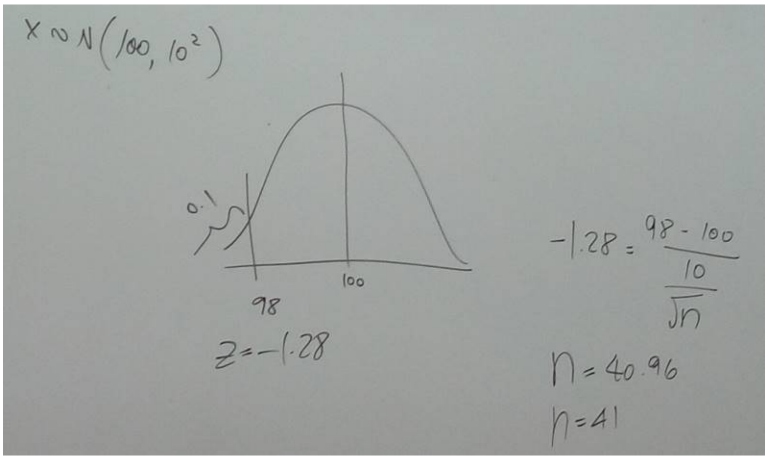




2. מהו הממוצע שרק 16% ממוצעי המדגמים בגודל 100 יהיו קטנים ממנו?



3.ערכים מתפלגים נורמלית עם 100 ו 10בריבוע, מהו גודל המדגם שייתן בהסתברות של 0.1 ממוצעים שיהיו קטנים מ98



תמיד מעגלים כלפי מעלה!

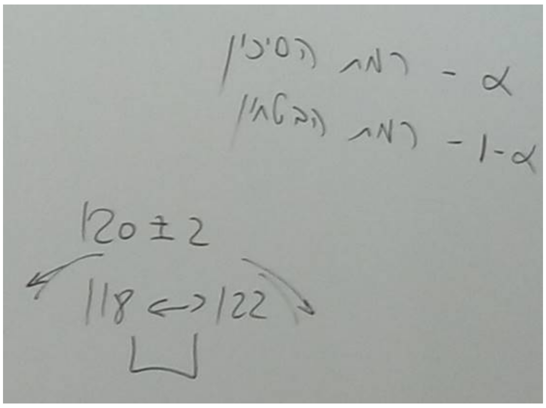
<u>הסקה סטטיסטית/ לאוכלוסייה.</u>

<u>החטיבות המרכזיות:,</u>

- **אמידה** - הפואנטה היא שהפרמטר של האוכלוסייה לא ידוע (התוכלת) ואנחנו על סמך תוצאות המדגם רוצים לאמוד מהי התוכלת
- **בדיקת השערות –** יש פרמטר ידוע, ואני מפעיל מניפולציה, האם כתוצאה מהמניפולציה הפרמטר השתנה באופן כזה או אחר?

<u>אמידה -</u> מהו אומדן טוב?

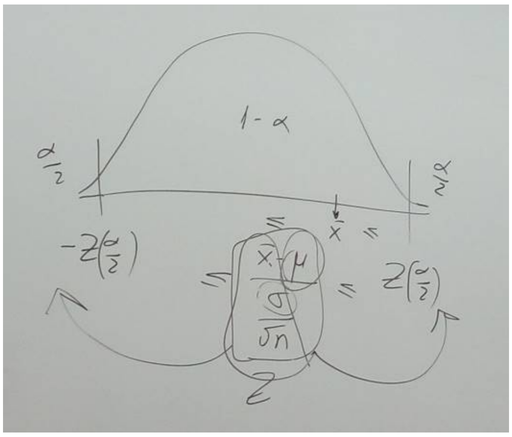
- האומדן חייב להיות **חסר הטיה** – בשביל זה המדגם שלי חייב להיות מקרי ומייצג באופן נורמלי את האוכלוסייה. אם אין מדגם מקרי ומייצג לא ניתן יהיה לתת אומדן!
- על האומדן להיות **יעיל** – ככל שהמדגם גדול יותר הפיזור קטן יותר והיעילות של האומדן תעלה, לכן תמיד נשאף למקסימום מדגם כמה שניתן בהתאם למשאבים העומדים לרשותו של החוקר, - לא מחייב – רצוי
- אומדן נקודתי – על סמך תוצאה אחת אני רוצה להסיק מהי התוכלת!



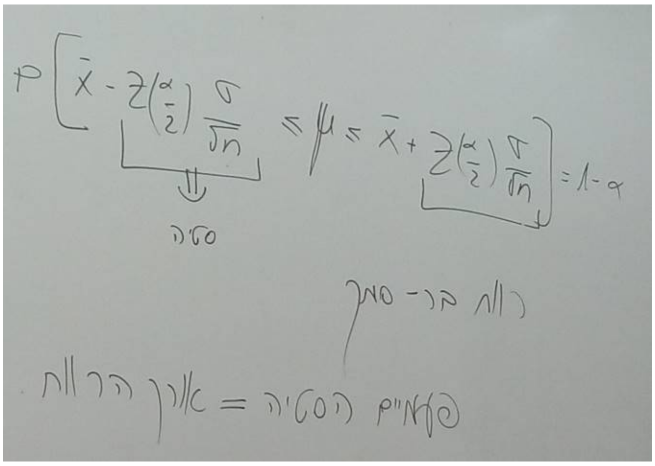
אלפא – אומר לי את רמת הסיכון

1- אלפא רמת הביטחון

כלל מפתח להיות בלפחות 95% ביטחון כלומר אלפא יהיה מקסימום 5%. הסיכון מחולק באופן סימטרי בשני צדי הטווח



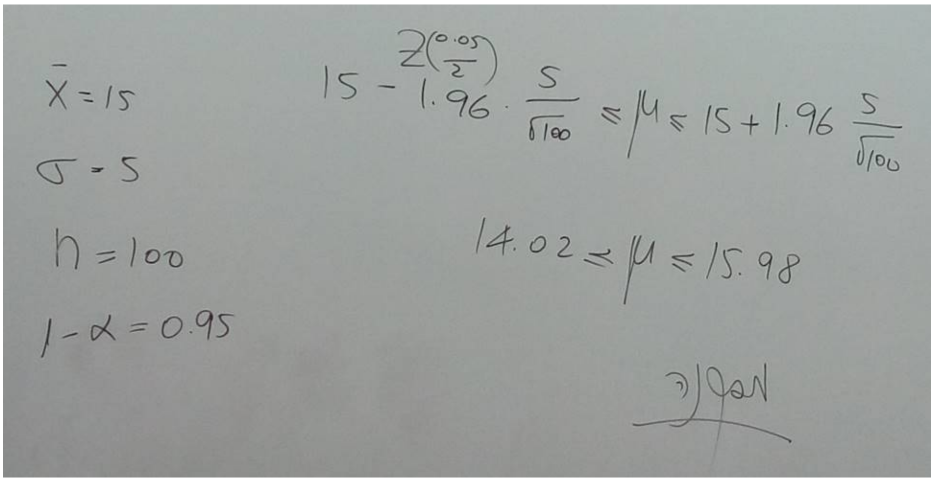
רווח בר סמך:



דוגמה:

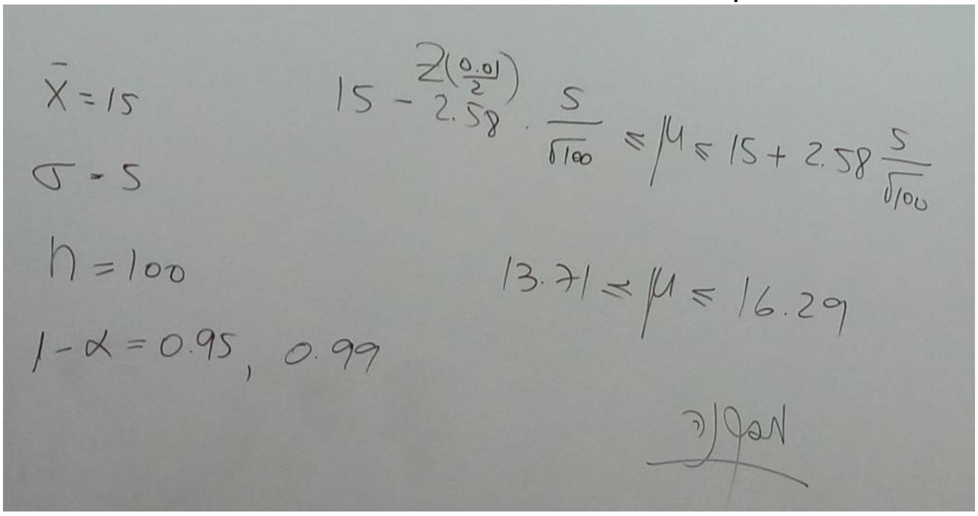
הממוצע – 15 סטיית התקן 5

בנה רווח בר סמך ברמת ביטחון של 95% לתוכלת האוכלוסייה עבור ממוצעים בגודל 100 (N= 100)

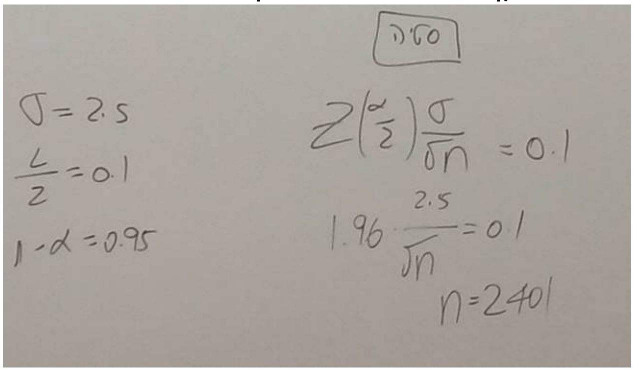


מתוך אינסוף דגימות חוזרות של מדגמים בגודל 100 (N בהתאם לכל שאלה) תוכלת האוכלוסייה תימצא ברווח המקרי שנע בין 14.02 ל15.98 ברמת ביטחון של 95%

אם נגדיל את הביטחון לרמה של 99% אז הרווח יגדל



מהו גודל המדגם הדרוש על מנת לאמוד את תוכלת האוכלוסייה, כאשר הסטייה היא 0.1 סטיית התקן = 2.5 ורמת הביטחון 95% ?



כדי לעמוד בדרישות אלו צריך 2401 נבדקים, אם אין כסף לזה אז אפשר להתפשר או על הסטייה או על רמת הביטחון, ואז יהיה אפשר להסתפק בפחות נבדקים.

<u>בדיקת השערות</u>