שיטות מחקר כמותניות



תוכן העניינים

1. מושגי יסוד בשיטות מחקר 1
2. מתאם וסיבתיות 15
3. דגימה 20
4. איכות כלי מדידה 27

שיטות מחקר כמותניות

פרק 1 - מושגי יסוד בשיטות מחקר

תוכן העניינים

1. מושגי יסוד בשיטות מחקר 1

מושגי יסוד בשיטות מחקר:

הגישה המדעית:

רקע:

פרק זה נעמוד על מאפייני השיטה המדעית, נגדיר מהי תיאוריה מדעית, מהי בעיית מחקר והשערת מחקר. נלמד בנוסף על סוגים של משתנים ועל האופן שבו הם קובעים את סוג המחקר.

הגישה המדעית - כיצד היא נבדלת מגישות אחרות לאיסוף ולעיבוד מידע על העולם?

שיטות שונות לרכישת ידע (צ'ארלס פירס):

1. שיטת הדבקות: האדם דבק בדעותיו ללא הטלת ספק או בחינת מידע אחר.
2. שיטת הסמכות: האדם מקבל את המידע על ידי הסתמכות על גורם אחר הנתפס כבעל ידע .
3. השיטה האינטואיטיבית: האדם יקבל תשובה לשאלה אם זו נראית לו הגיונית ומתוך תחושה שזו החלטה טובה .
4. הגישה הרציונאלית: בגישה זו הידע מושג באמצעות חשיבה והסקת מסקנות. תהליך זה נקרא הֶ קֵ ש.
5. אמפריציזם: ידע מבוסס בראש ובראשונה על הניסיון - תצפית ומדידה.
6. הגישה המדעית: שילוב של הגישה הרציונאלית והאמפירית. איסוף שיטתי של נתונים ברי מדידה והסקה לוגית של מסקנות מתוך הממצאים.

הגישה המדעית נבדלת מקודמותיה במתודולוגיה (דרך החקירה) ולא באמיתות מסקנותיה.

העקרונות המנחים אותה:

1. אובייקטיביות - משוחררת משיפוטים ערכיים והעדפות סובייקטיביות.
2. אמפיריות - מושא המחקר ניתן לחקירה ניסויית (אמפירית). מבוססת על נתונים ברי מדידה.
3. סקרנות - הכוח המניע של המחקר המדעי.
4. ספקנות - הטלת ספק בכל טענה או גילוי עד אשר הם יעמדו למבחן מדעי.
5. כלליות - המטרה של המדע היא לנסח חוקים כלליים.

תפקיד המדע הוא לגבש ולארגן חוקים והסברים כלליים (כלומר שחלים על מגוון רחב של מצבים, דרכי תגובה ואנשים ולא על אירוע מסוים אחד) במטרה להבין ולהסביר את העבר, לנבא ולשלוט בעתיד.

תרגול:

1) משה התווכח עם החבר שלו שמעון לגבי החשיבות של הפסקת עישון. משה הביא עדויות של רופאים רבים שטוענים שהעישון מזיק לבריאות ואילו שמעון הגיב – לא מעניין אותי! אני חושב שעישון לא מזיק וזהו.
השלם את המשפט: "משה מדגים שיטת ה_____ של רכישת ידע ואילו שמעון מדגים את שיטת ה_____".

א. סמכות, דבקות.
ב. מדעית, דבקות.
ג. מדעית, אינטואיטיבית.
ד. סמכות, אינטואיטיבית.

2) מהם חמשת המאפיינים לגישה המדעית?

א. השערה, עניין, סובייקטיביות, ידע רב, פרטניות.
ב. סקרנות, מהימנות, כלליות, אימפריות, אובייקטיביות.
ג. סקרנות, ספקנות, אובייקטיביות, כלליות, אימפריות.
ד. אף תשובה אינה נכונה.

3) איזה מהטענות הבאות מהווה טענה מדעית?

א. הנשמה ממשיכה להתקיים אחרי שהגוף מת.
ב. לבעלי חיים צריכים להינתן זכויות בסיסיות הניתנות לבני האדם.
ג. יונקים הם חיות חמודות יותר מזוחלים.
ד. רוב האנשים יצייתו לגורם סמכות גם אם הדבר נוגד את ערכיהם.

תשובות סופיות:

1) א' **2)** ג' **3)** ד'

סולמות מדידה:

רקע:

משתנה = מדד היכול לקבל שני ערכים לפחות.

למשל:

משתנה	ערכיו
רמת תוקפנות	1 עד 10
צבע עיניים	חום, ירוק, כחול או שחור
מגדר	זכר, נקבה
סטאטוס בעבודה	פועל, מנהל משמרת, מנהל סניף, מנהל כללי

בניגוד למשתנה, קבוע מוגדר כמדד היכול לקבל ערך אחד בלבד.
למשל:

קבוע	ערכיו
רמת תוקפנות	5
מגדר	זכרים בלבד

מה שמבדיל בין משתנה לקבוע זה השונות.

משתנה: $\sigma^2 > 0$.

קבוע: $\sigma^2 = 0$.

סולם מדידה = הסולם שבו נמדדו ערכיו של המשתנה. הסולם קובע את אופיו של המשתנה כאיכותי או ככמותי ואת סוג המדדים הסטטיסטיים שניתן לחשב.

שמי – מאפשר לחוקר להסיק על זהות המשתנה בלבד (כגון: צבע עיניים, מגדר, מוצא, אזור מגורים). אין משמעות למספרים. ניתן לחשב שכיח בלבד.

סדר – סולם גבוה יותר, שבו יש חשיבות גם לסדר בין הערכים ולא רק לזהות. ניתן לדרג מהנמוך לגבוה (כגון: דרגות בצה"ל, מקום זכייה בתחרות, רמת השכלה וכו'). ניתן לחשב גם מדדי מיקום מרכזי כמו חציון ומדדי שונות של טווח בין רבעוני.

משתנים בסולם מדידה שמי או סדר מוגדרים כמשתנים איכותיים.

רווח – סולם גבוה יותר, שבו מלבד לסדר יש חשיבות גם להפרשים בין הערכים. ניתן לדעת בכמה ערך אחד גדול מהשני (כגון: טמפ', ציון IQ של ווקסלר, ציון פסיכומטרי וכו'). הערך 0 הוא ערך שרירותי / יחסי שלא מייצג היעדר תכונה. ניתן לבצע את כל הניתוחים הסטטיסטיים (כמו חישוב ממוצע וסטיית תקן/שונות).

מנה – הסולם הגבוה ביותר שמאפשר להסיק גם על יחסים. כלומר, פי כמה ערך אחד גדול מהשני ולא רק בכמה. הסקה על יחסים מתאפשרת בעקבות כך שהערך 0 הוא מוחלט ומייצג במשמעות היעדר תכונה (כגון: גובה, משקל, משך זמן, ציון מבחן וכו').

משתנים בסולם רווח או מנה מוגדרים כמשתנים כמותיים.

טרנספורמציות על סולמות המדידה:
ביצוע שינוי מגמתי על מערכת המספרים, בתנאי שהסולם משמר את התכונות שלו.
טרנספורמציה ליניארית – הכפלה/הוספה של קבוע ו/או הכפלה/חילוק בקבוע.
טרנספורמציה שאיננה ליניארית – כמו העלאה בריבוע או הוצאת שורש.

ככל שהסולם גבוה יותר כך ניתן לבצע עליו פחות טרנספורמציות:

סולם שמי: כל טרנספורמציה שומרת זהות – אסור להכפיל ב-0.

סולם סדר: כל טרנספורמציה שומרת סדר (וזהות).
פרט לטרנספורמציות האסורות בסולם שמי, בסולם סדר אסור גם להכפיל במספר שלילי.

סולם רווח: טרנספורמציה השומרת על היחס בין ההפרשים (וגם על הסדר והזהות).
טרנספורמציה ליניארית חיובית (הוספה/החסרה של קבוע והכפלה בקבוע חיובי בלבד).

סולם מנה: טרנספורמציה שומרת יחס (כמו גם על הזהות, הסדר וההפרשים).
הכפלה בקבוע חיובי בלבד.

סיכום:

הסולם	מה מלמד הסולם?	טרנספורמציות מותרות
סולם שמי – Nominal scale	זהות	טרנספורמציה שומרת זהות (אסור להכפיל באפס)
סולם סדר – Ordinal scale	זהות, סדר	טרנספורמציה שומרת סדר (אסור להכפיל בקבוע שלילי)
סולם רווח – Interval scale	זהות, סדר, הפרשים	טרנספורמציה ליניארית חיובית
סולם מנה – Ratio scale	זהות, סדר, הפרשים, יחסים	הכפלה בקבוע חיובי בלבד

תרגול:

1) עבור כל אחד מהמשתנים הבאים – רשמו מהו סולם המדידה ונמקו:

א. רמת השכלה: יסודית, תיכון, אקדמאית.

ב. כדי לבדוק את הערך התזונתי של מזון תינוקות חדש נשקלו 100 תינוקות.

ג. התפלגות הסטודנטים לפי אזור מגורים: צפון מרכז ודרום.

ד. "זמן תגובה" של נהגים בשניות.

ה. דרגות שכר: מעל לממוצע, מתחת לממוצע, ממוצע .

ו. טמפרטורת התינוקות הנמדדת בטיפת חלב.

ז. תוצאות המבחנים בקורס שיטות מחקר.

ח. השעה בה משכים אדם מדי בוקר, במשך חודש ימים.

ט. פסיכולוג בדק רמת חרדה שדווחה על סקאלה בת 7 רמות: 1-7.

י. שנת הלידה של נבדקים וגילם.

יא. הגדרה דתית: חרדי, דתי, מסורתי, חילוני.

2) חוקר קידד את המצב המשפחתי באופן הבא: 1 – רווקים, 2 – נשואים, 3 – גרושים. אילו מהטרנספורמציות הבאות אסור לו לבצע?

א. 4,3,2.

ב. 6,4,1.

ג. -1,-2,-3.

ד. 1,2,1.

ה. 9,4,1.

3) חוקר קידד משתנה של דרגות בצבא באופן הבא: 1 – טוראי, 2 – רב"ט, 3 – סמל. אילו מהטרנספורמציות בסעיף הקודם אסור לו לבצע?

א. 4,3,2.

ב. 6,4,1.

ג. -1,-2,-3.

ד. 1,2,1.

ה. 9,4,1.

4) חוקר אסף נתונים על הטמפ' במהלך כמה ימים בחודש נובמבר וקיבל את התוצאות הבאות: 20, 22, 30. אילו מהטרנספורמציות הבאות מותר לו לבצע על הנתונים?
   א. $-30,-22,-20$.
   ב. $1,2,1$.
   ג. $21,20,19$.
   ד. $62,46,42$.
   ה. $60,44,40$.

5) בהנחה שהמשתנה הנמדד הוא כמות הגשם שירד בנובמבר, אילו מהטרנספורמציות של הסעיף הקודם מותרות?

תשובות סופיות:

1) א. איכותי אורדינאלי. ב. כמותי מנתי. ג. איכותי נומינאלי.
   ד. כמותי מנתי. ה. איכותי אורדינאלי. ו. כמותי אינטרוואלי.
   ז. כמותי מנתי. ח. כמותי אינטרוואלי. ט. כמותי אינטרוואלי.
   י. שנת לידה: כמותי אינטרוואלי, גיל: כמותי מנתי.
   יא. איכותי נומינאלי.
2) ד'.
3) ג' ו-ד'.
4) ד' ו-ה.
5) ה.

תיאוריה מדעית:

רקע:

הגדרה: "תיאוריה היא מערכת של מונחים, הגדרות וטענות המייצגות השקפה על תופעה מסוימת על-ידי הצגת היחסים בין המונחים במטרה להסביר ולנבא את התופעה" (קרלינגר, 1972).

להגדרה שלושה מרכיבים:
א. התיאוריה כוללת מערכת טענות על קשרים בין מונחים
ב. התיאוריה מציגה השקפה מסוימת על התופעה
ג. לתיאוריה יש תפקיד של הסבר וניבוי

תיאוריה טובה צריכה להיות:

1. מופשטת: בסיס מושגי אבסטרקטי אשר מנותק לכאורה מן המציאות הממשית שבה היא דנה.
2. כוללנית: תיאוריה אמורה להסביר חוקיות או דפוס שחוזר על עצמו. להסביר את מקסימום המקרים הרלוונטיים.
3. חסכונית: ההסבר צריך להיות יפה, אסתטי וכולל מספר משתנים מצומצם שיסבירו כמה שיותר.
4. תקפה לוגית: היגיון פנימי בקשר שבין ההנחות לבין ההסברים והתוצאות.
5. בהירה ופשוטה: תיאוריה טובה מנוסחת בבהירות כך שניתן להבין את ההיגיון ואת הטענות שלה.
6. בעלת יכולת אישוש והפרכה: קבלה אם העדויות האמפיריות מאמתות אותה, ודחייה אם העדויות האמפיריות מפריכות אותה.

קריטריון ההפרכה = הקריטריון המרכזי שמבחין בין תיאוריה מדעית לפסאודו-מדעית (פופר, 1962). תיאוריה שאין דרך אמפירית לנסות להפריכה איננה תיאוריה מדעית.

בהקשר זה טבע פופר את המונח – "**ניבוי מסתכן**". ניבוי שאם הוא לא מתקיים, אזי התיאוריה מופרכת.

כאשר ההשערה הנגזרת מהתיאוריה מופרכת ניתן להסיק כי:

1. המחקר לא היה טוב.
2. התיאוריה לא נכונה.
3. פתרון ביניים – שינוי או עדכון התיאוריה.

בעיית / שאלת מחקר:

בעיה או שאלה ייחודית שניתן לחקור אותה בכלים מדעיים.

למשל:
תחום מחקר - "השתלבות בצבא".
שאלת מחקר - "האם טיב התפקוד בצבא קשור לסוג היישוב?".

בעיית המחקר מנוסחת תמיד בצורת שאלה וצריכה לעמוד בארבעה קריטריונים כדי שיהיה אפשר לחקור אותה בשיטה המדעית:

1. הבעיה צריכה לבטא **יחס בין שני משתנים** או יותר.
2. הבעיה חייבת להיות **מנוסחת בצורה ברורה.**
3. הבעיה חייבת **לעמוד בבחינה אמפירית**. כלומר, צריכה להיות דרך מעשית למדוד את המשתנים המוצגים בשאלת המחקר.
4. **אובייקטיבית**. לא מבטאת את ציפיות החוקר.

משאלה להשערה:

השערת המחקר היא ממוקדת וספציפית יותר משאלת המחקר . בדרך כלל בעלת כיוון ומבטאת את ציפיות החוקר באשר לתוצאות המחקר. היא מנוסחת לפני איסוף הנתונים כאשר הנתונים שיאספו יאששו או יפריכו אותה.

למשל:
שאלת מחקר – "האם טיב התפקוד בצבא קשור לסוג היישוב?".
השערת מחקר – "בני קיבוצים מתפקדים טוב יותר מאחרים בצבא".

בעוד שבעיית המחקר מנוסחת בצורת שאלה, השערת המחקר **מנוסחת בצורת קביעה** ומתייחסת **לאופי הקשר בין המשתנים**.

סוגי השערות:
השערה עם כיוון ("בני קיבוצים מתפקדים טוב יותר מאחרים בצבא").
השערה ללא כיוון ("טיב התפקוד של בני קיבוצים בצבא שונה מאחרים").
היעדר קשר ("אין הבדל בתפקוד בצבא בין בני קיבוצים לאחרים").

סיכום – משאלה להשערה:

בעיה כללית	שאלת מחקר	השערת מחקר
ניסוח תחום המחקר (כללי). למשל:	מנוסחת בצורת שאלה. למשל:	מנוסחת בצורת קביעה. למשל:
"יש סטודנטים עם יכולת גבוה שלא מצליחים להגיע להישגים גבוהים"	"האם דימוי עצמי משפיע על ההישג?"	"סטודנטים עם דימוי עצמי נמוך יגיעו להישגים נמוכים יותר מאשר סטודנטים עם דימוי עצמי גבוה."

תרגול:

1) נסחו השערות מחקר לשאלות הבאות:

א. האם הלימוד בקבוצה משפר את ההישגים?

ב. האם חוסר איזון בין הקוגניציות גורם לדיסוננס?

ג. האם שימוש במריחואנה פוגע בזיכרון לטווח קצר?

ד. האם קיים קשר בין השכלת ההורים למידת ההצלחה האקדמית של ילדיהם?

תשובות סופיות:

1) ראו סרטון.

השערות מחקר:

רקע:

קביעה העוסקת בקשר בין 2 משתנים לפחות.

על מנת שטענה / השערה כלשהי תחשב הסבר מדעי היא צריכה לקיים 2 דרישות:

1. דרישת הבחינות – ההשערה חייבת להיות ניתנת לבדיקה אמפירית. כלומר שהמשתנים בהשערה יהיו ניתנים לתצפית ולמדידה כך שניתן יהיה לכמת אותם.
2. דרישת הרלוונטיות – ההשערה חייבת להיות מנוסחת כך שניתן יהיה לחשב מתאם סטטיסטי. במילים אחרות, ההשערה צריכה להיות מנוסחת כקשר בין משתנים ולא בין קבועים כדי שניתן יהיה לחשב מתאם.

טענות / השערות מנוסחות כקשר בין שני סוגים של משתנים:
משתנה בלתי-תלוי (מב"ת): משתנה המסביר, המשפיע, כלומר הסיבה.
משתנה תלוי (מ"ת): משתנה המוסבר, המושפע, כלומר התוצאה.

תרגול:

1) לפניכם רשימת טענות. לגבי כל טענה רשמו:

1. מהו המשתנה הב"ת (וערכיו)?
2. מהו המשתנה התלוי (וערכיו)?
3. האם הטענה עומדת בדרישת הרלוונטיות?
4. האם הטענה עומדת בדרישת הבחינות?
5. האם הטענה מהווה הסבר מדעי?

א. אנשים דומים נמשכים אחד לשני יותר מאשר אנשים שונים.
ב. ילדים להורים בעלי השכלה אקדמאית הם בעלי אינטליגנציה גבוהה.
ג. היצורים בכוכב הלכת נגה הם ירוקים.
ד. בתנאי לחץ גבוה מתקבלות החלטות באיכות נמוכה יותר מאשר בתנאי לחץ נמוך.
ה. אימהות מבלות שעות רבות יותר עם ילדיהן לעומת האבות, בתנאי שמדובר בבני זוג נשואים.
ו. בני מזל בתולה הם קפדנים.
ז. בני האדם אינטליגנטים יותר מאשר חוצנים.
ח. ככל שהחום עולה כך אנשים נוטים להתרגז יותר.

ט. אנשים בעלי השכלה נמוכה מאמינים באמונות טפלות יותר מאשר אנשים בעלי השכלה בינונית וגבוהה.

י. מספר תאונות הדרכים בקרב נהגים צעירים גבוה יותר מאשר בקרב נהגים ותיקים.

תשובות סופיות:

1) ראו סרטון.

מערכי מחקר:

רקע:

לכל אחד מהמשתנים נבחין בין שתי רמות:

1. רמה תיאורטית / נומינאלית – השם המילולי של המשתנה שרוצים לחקור.
2. רמה תצפיתית / אופרציונלית – הכוונה לאופן שבו בדקו / מדדו את המשתנה באותו מחקר. במילים אחרות, כיצד הפכו את המשתנה למדד כמותי.

המשתנים התיאורטיים הם פעמים רבות מונחים מופשטים ולכן קשה להגדיר אותם תצפיתית.

כשהמעבר בין הרמה התיאורטית לרמה התצפיתית אינו חד משמעי יתכנו 2 בעיות:

1. חוסר מיצוי – התצפית לא מבטאת את כל ההיבטים של המשתנה התיאורטי.
2. חוסר בלבדיות – המשתנה התצפיתי מכיל משתנים אחרים מעבר למשתנה התיאורטי.

אנחנו מבחינים בין שני סוגים של הגדרות אופרציונליות:

1. הגדרה אופרציונלית נמדדת – הגדרה המתארת את האופן שבו החוקר מודד את הערכים השונים של המשתנה. כאשר לא קיימת מניפולציה מצד החוקר אלה החוקר מודד מצב נתון שהנבדקים הגיעו אתו למחקר.
2. הגדרה אופרציונלית ניסויית – הגדרה המתארת את האופן שבו החוקר יוצר את הערכים השונים של המשתנה. כלומר, זוהי הגדרה המתארת מניפולציה בה החוקר מתפעל ערכים שונים של המשתנה ולא מודד מצב נתון.

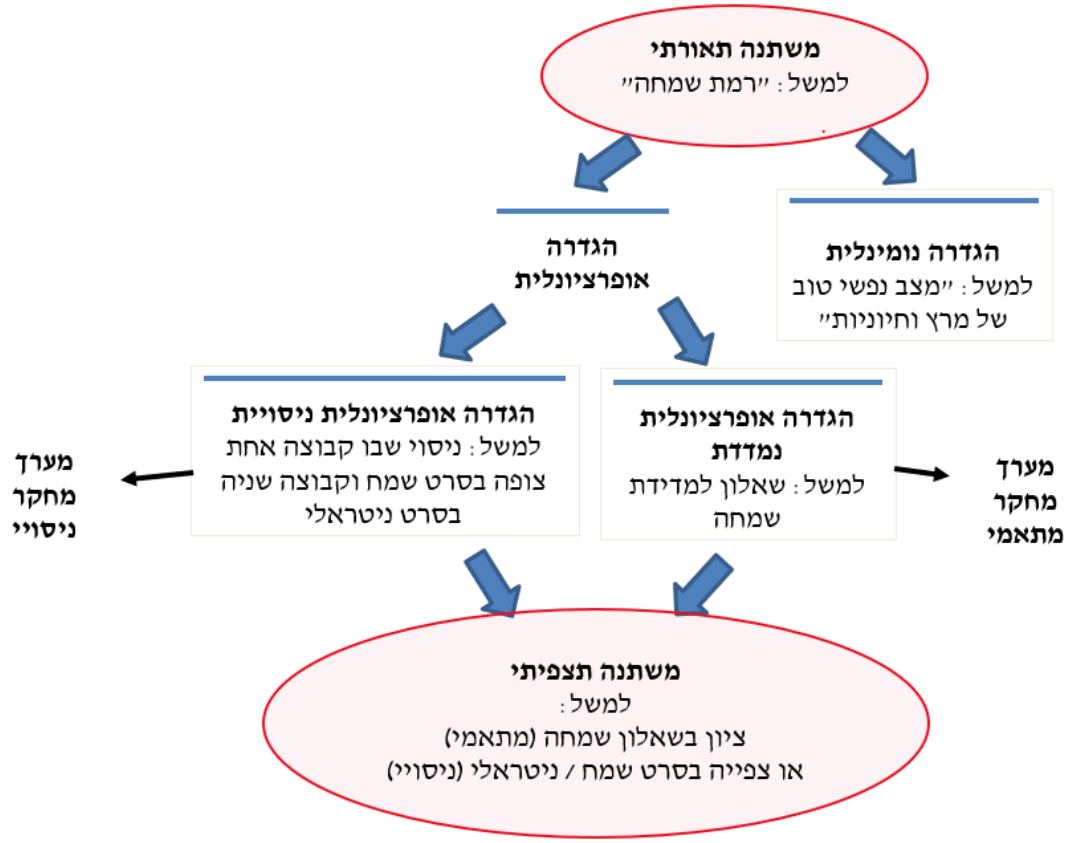
מערך מחקר מתאמי לעומת ניסויי:
סוג ההגדרה האופרציונלית שבה החוקר משתמש למשתנה הב"ת תגדיר את סוג מערך המחקר שבו נקט החוקר:
הגדרה אופרציונלית נמדדת למשתנה הב"ת ⇐ מערך מחקר מתאמי.
הגדרה אופרציונלית ניסויית למשתנה הב"ת ⇐ מערך מחקר ניסויי.

ההבדל בין מערך מחקר מתאמי לניסויי:
מערך מחקר מתאמי – החוקר מודד 2 משתנים ולכן יכול להסיק אך ורק על קיום / אי קיום של קשר סטטיסטי ביניהם.
ככל ש-A עולה/יורד כך B עולה/יורד ולהפך.

מערך מחקר ניסויי – החוקר מתפעל את המשתנה הב"ת ולכן יכול להסיק בנוסף לקיום קשר גם על קיום סיבתיות. כלומר, יכול להסיק נכון את הכיוון של הקשר ולהפריך הסברים אחרים לקשר פרט למשתנה הב"ת.
A משפיע על B.

משתנה ייחוס – משתנה שהחוקר לא יכול לתפעל (ליצור) את הערכים שלו, אלא אך ורק למדוד מצב נתון שלהם כגון: גיל, מין, גובה, מוצא צבע עיניים.
כאשר המשתנה הב"ת במחקר הוא משתנה ייחוס (ניתן למדידה בלבד) אזי החוקר יכול לערוך רק מערך מחקר מתאמי.

סיכום – ממשתנה תאורתי לתצפיתי:



תרגול:

1) לפניך שני תיאורים של מחקרים:
תיאור א': חוקר רצה לבדוק האם ג'ינג'ים הם חמי מזג. הוא דגם 40 אנשים בעלי שיער ג'ינג'י ונתן להם שאלון שבודק את סף התסכול שלהם (1-5).
תיאור ב': חוקר רצה לבדוק האם רמת המוטיבציה משפיעה על איכות ביצוע מטלות. הוא דגם 60 אנשים והבטיח להם כרטיס טיסה לחו"ל ליעד נחשק כפרס על ביצוע טוב של המטלה. הוא נתן להם מבחן בחשבון (ציונים 0-10).
עבור כל תיאור ציינו:

א. מהם המשתנים התלוי והב"ת התיאורטיים והתצפיתיים?

ב. מהו מערך המחקר בו נקט החוקר – מתאמי או ניסויי?

ג. האם המחקר עומד בדרישות להסבר מדעי (הבחינות והרלוונטיות)?

ד. במידה והמחקר לא עומד בדרישות, כיצד ניתן לתקנו?

תשובות סופיות:

1) א. תיאור א': תיאורטיים – תלוי: מידת היות האדם חג מזג, ב"ת: צבע שיער.
תצפיתיים – תלוי: ציון בשאלון סף תסכול, ב"ת: צבע שיער.
תיאור ב': תיאורטיים – תלוי: איכות ביצוע מטלה, ב"ת: רמת המוטיביציה.
תצפיתיים – תלוי: מס' תשובות נכונות במבחן חשבון, ב"ת: הבטחה של כרטיס לחו"ל תמורת ביצוע טוב של המטלה.
ב. תיאור א': מתאמי, תיאור ב': ניסויי.
ג. לא. ד. ראו סרטון.

שיטות מחקר כמותניות

פרק 2 - מתאם וסיבתיות

תוכן העניינים

1. מתאם וסיבתיות 15

מתאם וסיבתיות:

סכמות לתיאור קשרים:

רקע:

מתאם סטטיסטי בין שני משתנים משמעו הוא שככל שמשתנה אחד עולה / יורד כך המשתנה השני עולה / יורד ולהיפך.
למשל: ככל שההשכלה עולה כך המעמד הסוציו אקונומי גבוה יותר או להיפך.

לא ניתן לדעת מה גרם למה – ההשכלה לעליה במעמד או שמא להיפך מעמד מאפשר השכלה גבוהה יותר. במילים אחרות, לא ניתן לדעת מה קדם למה – השכלה למעמד או להפך – מעמד להשכלה. בנוסף יתכנו הסברים חלופיים לקשר שבין השכלה למעמד, כגון רווחה כלכלית.

סיבתיות ניתנת להסקה רק כאשר מתקיימים 3 התנאים הבאים:

1. ימצא קשר סטטיסטי / מתאם בין שני המשתנים (B,A)
   תנאי הכרחי, אך לא מספיק לצורך הסקה סיבתית.
2. שליטה על סדר זמנים = קדימות בזמן
   המב"ת (הסיבה) קדם בזמן למ"ת (התוצאה).
3. הפרכת הסברים חלופיים
   היכולת להוכיח שמה שהשפיע על המ"ת זה רק המב"ת ולא משתנים אחרים.

במחקר ניסויי:
החוקר מחלק את הנבדקים באופן מקרי לקבוצות שונות ולכן יכול להניח שבתחילת הניסוי קיים שוויון התחלתי במאפייני הנבדקים במשתנים שונים.
החוקר מתפעל את המב"ת ואז מודד את המשתנה התלוי.
אם בתום הניסוי ימצא הבדל בין הקבוצות:

1. קיים קשר סטטיסטי בין המב"ת למ"ת – נמצא הבדל בין הקבוצות.
2. קיימת שליטה בסדר הזמנים - ההבדל בין הקבוצות נוצר רק לאחר הפעלת המניפולציה.
3. הופרכו הסברים חלופיים - ניתן להסיק שלא קיימים משתנים אחרים שיכולים להסביר את התוצאה מלבד המניפולציה כי הקבוצות היו שוות במאפייני הנבדקים טרם הפעלת המניפולציה.

שלושת התנאים להסקת סיבתיות מתקיימים במערך המחקר הניסויי.

במחקר מתאמי:
החלוקה לקבוצות היא לא מקרית כי נערכת בהתאם לערך שהנבדק הגיע עימו למחקר או שהסיווג לקבוצות היה בהתאם לדיווח של הנבדקים.

אם בתום המחקר ימצא קשר סטטיסטי או הבדל בין הקבוצות:

1. קיים קשר סטטיסטי בין המב"ת למ"ת.
2. לא ניתן להוכיח שהמב"ת קדם בזמן לתלוי – יכול להיות שהתלוי השפיע על הב"ת ולא להיפך.
3. לא ניתן לדעת שההבדל היחידי היה במב"ת וייתכנו הסברים חלופיים לקשר שנמצא.

<u>רק התנאי הראשון לסיבתיות יכול להתקיים במערך מחקר מתאמי ולכן לא ניתן להסיק על סיבתיות.</u>

הסיבות לשימוש במערך מתאמי:

1. יש משתנים <u>שלא ניתן</u> לעשות בהם מניפולציה (אינטליגנציה, מגדר וכו').
2. ישנם משתנים שבהם <u>לא מוסרי</u> לעשות מניפולציה (תוקפנות, רמת חרדה וכו').
3. המערך המתאמי <u>פחות מלאכותי</u> וקרוב יותר למציאות (מחקרי שדה ולא מעבדה).

סכמות לתיאור קשר סיבתי:

א. קשר ישיר $(A \Leftarrow B)$.

ב. קשר הפוך $(A \Rightarrow B)$.

ג. קשר מזויף (C ⇐ A, C ⇐ B) - ולכן לא קיים קשר אמיתי בין A ל-B.

ד. קשר מתווך (A ⇐ C ⇐ B).

ה. קשר מותנה (A ⇐ B בתנאי * C).

ו. קשר ממותן (A ⇐ B שונה ברמות שונות של ** C).

* C הוא קבוע

**C הוא משתנה איכותי

סיכום – סוגי קשרים סיבתיים:

סוג הקשר	סכמת הקשר	תיאור מילולי	C מכונה	C מהות
מזויף	A ← C → B	C קשור גם ל-A וגם ל-B	מתערב	כמותי / איכותי
מתווך	A → C → B	A גורם ל-C שגורם ל-B	מתווך	כמותי
מותנה	A → B (בתנאי C)	A גורם ל-B רק בתנאי C	תנאי	קבוע
ממותן	A → B (שונה ברמות השונות של C)	הקשר בין A ל-B חזק / חלש יותר עבור ערך מסוים של C	ממתן	איכותי

שימו לב – כל הקשרים חוץ ממתאם מזויף, מתארים קשר סיבתי.

תרגול:

1) במחקר נמצא קשר בין רמת מעורבות חברתית (A) לבין רמת הישגים לימודיים (B). במטרה להסביר את המתאם הוצעו הטענות הבאות. הציגו את הקשר הסיבתי בין המשתנים המופיעים בכל טענה באופן סכמטי, וציינו במילים מה מבטאות האותיות השונות.

א. ככל שהאדם מופנם יותר כך מעורבותו החברתית נמוכה יותר וגם רמת הישגיו בלימודים נמוכה יותר.

ב. רמת ההשכלה של ההורים קשורה למידת המעורבות החברתית וגם לרמת ההישגים בלימודים של הילד.

ג. בכיתות נמוכות בלבד, ככל שרמת ההישגים עולה כך מידת המקובלות החברתית עולה, מה שתורם למידת המעורבות החברתית.

ד. תלמידים מעורבים יותר חברתית נוטים ללמוד בקבוצות לימוד באופן שתורם לעליה בהישגיהם בלימודים.

ה. מעמד חברתי כלכלי קשור לרמת ההישגים בלימודים וגם למידת המופנמות שמשפיעה בתורה על רמת המעורבות החברתית.

ו. המעמד החברתי כלכלי משפיע על מידת המופנמות שמשפיעה הן על מידת המעורבות החברתית והן על ההישגים בלימודים.

2) שלושה חוקרים רצו לבדוק את הטענה כי מצב הרוח משפיע על מידת המוכנות להושיט עזרה לזולת. שלושת החוקרים השמיעו לנבדקים הקלטה של אדם המבקש עזרה ומדדו את הזמן שלקח להם לפתוח את הדלת מרגע שמיעת ההקלטה.
החוקר הראשון – חילק לנבדקים שאלון מצב רוח (1-רע עד 15-מצוין). החוקר מצא קשר חיובי בין הציון בשאלון לבין הזמן שלקח לנבדקים לפתוח את הדלת.
החוקר השני – חילק אף הוא לנבדקים את השאלון למדידת מצב הרוח ואח"כ חילק אותם לשלוש קבוצות לפי הציון בשאלון: רע=ציון 1 עד 5, ניטראלי=6 עד 10 וטוב=11 עד 15. החוקר מצא הבדל מובהק בין שלושת הקבוצות בזמן שלקח לנבדקים לפתוח את הדלת.
החוקר השלישי – השמיע לנבדקים סיפורים שונים הנבדלים במצב הרוח שמשרים: סיפור שמח, סיפור נטרלי וסיפור עצוב. החוקר לא מצא הבדל בין שלושת הקבוצות בזמן שלקח לנבדקים לפתוח את הדלת.

א. מהם המשתנים התיאורטיים של כל אחד משלושת החוקרים?

ב. מהם המשתנים התצפיתיים של כל אחד משלושת החוקרים?

ג. מהו מערך המחקר שבו נקט כל אחד מהחוקרים?

ד. מה משמעות הממצאים של חוקר ב'? האם הוא יכול להסיק על סיבתיות?

ה. האם חוקר ג' יכול להסיק על סיבתיות?

ו. חוקר ד' טען כי מידת החברותיות משפיעה הן על מצב הרוח והן על מידת הנכונות לעזרה לזולת. כיצד יש בטענתו בכדי לערער על ממצאי המחקר של שני החוקרים הראשונים?

ז. כיצד חוקרים א' ו-ב' יכולים להגן על ממצאיהם בפני טענתו של חוקר ד'?

3) במחקר נמצא קשר בין משך זמן נטילת תרופות נוגדות דיכאון (A) לבין עליה במשקל (B).
נמצא כי ככל שאדם נוטל תרופות אלו למשך זמן ממושך יותר כך ההפרש בין משקלו הנוכחי למשקלו ההתחלתי (לפני שהחל ליטול את הכדורים) גבוה יותר.

א. ידוע כי אחת מתופעות הלוואי של תרופות נוגדות דיכאון היא תיאבון מוגבר. כיצד ממצא זה עשוי להסביר את הקשר בין נטילת התרופה לעליה במשקל? הציגו את הקשר באופן סכמטי.

ב. נמצא בנוסף כי הקשר בין משך זמן נטילת התרופה לבין עליה במשקל חלש יותר עבור אנשים העוסקים בפעילות גופנית (מתונה ומעלה) לעומת אנשים שאינם מתעמלים כלל.
המשתנה – "עיסוק בפעילות גופנית" מהווה במקרה זה:

i. משתנה מתווך.
ii. משתנה תנאי.
iii. משתנה מתערב.
iv. משתנה ממתן.

ג. חוקר א' טוען כי הסיבה לעליה במשקל היא איננה נטילת התרופה אלא פשוט גיל הנבדקים. הגיל קשור הן למשך זמן נטילת התרופה והן לעליה במשקל.
   i. לטענתו, המשתנה - "גיל" הינו משתנה:
      א. מתווך.
      ב. ממתן.
      ג. תנאי.
      ד. מתערב.
   ii. כיצד ניתן לבדוק את טענתו?

ד. לטענת חוקר ב' הקשר בין משך זמן נטילת התרופה לבין עלייה במשקל לא מתקיים עבור גברים.
איזה מהטענות הבאות נכונה?
   i. לפי טענת חוקר ב', מגדר הוא משתנה מתערב בקשר שבין משך זמן נטילת התרופה לעליה במשקל.
   ii. לפי טענת חוקר ב', הקשר בין משך זמן נטילת התרופה לעלייה במשקל הוא קשר מזויף.
   iii. לפי טענת חוקר ב', מגדר הוא תנאי לקיום הקשר בין משך זמן נטילת התרופה לעליה במשקל.
   iv. לפי טענת חוקר ב', נשים הן תנאי לקיום הקשר שבין משך זמן נטילת התרופה לעליה במשקל.

תשובות סופיות:

1) א. $A \leftarrow C$. ב. $\begin{matrix} A \leftarrow C \\ B \swarrow \end{matrix}$. ג. $A \leftarrow D \leftarrow B$. ד. $B \leftarrow C \leftarrow A$.
ה. $\begin{matrix} A \leftarrow D \leftarrow C \\ B \swarrow \end{matrix}$. ו. $\begin{matrix} A \leftarrow D \leftarrow C \\ B \swarrow \end{matrix}$.

2) א. ב"ת – מצב רוח, תלוי – מידת מוכנות להושיט עזרה לזולת.
ב. ב"ת: חוקר א' וב' – ציון בשאלון מצב רוח, חוקר ג' – השמעת סיפורים שונים הנבדלים במצב הרוח שהם משרים.
תלוי: חוקר א', ב' וג' – משך הזמן מרגע שמיעת ההקלטה ועד לפתיחת הדלת.
ג. חוקר א': מתאמי, חוקר ב': מתאמי, חוקר ג': ניסויי.
ד. קשר סטטיסטי, לא.
ה. לא.
ו. קשר מזויף.
ז. ראו סרטון.

3) א. $B \leftarrow C \leftarrow A$. ב. iv. ג. i. ד'. ii. ראו סרטון. ד. iv.

שיטות מחקר כמותניות

פרק 3 - דגימה

תוכן העניינים

1. דגימה 20

דגימה:

מושגי יסוד בדגימה:

רקע:

השערות המחקר נוגעות תמיד לאוכלוסייה, אך מאילוצים שונים נאלץ לבצע את המחקר על חלק קטן ממנה – מדגם ונקיש מתוצאות המדגם על האוכלוסייה (היקש הסתברותי).
ככל שהמדגם ייצג / ישקף טוב יותר את האוכלוסייה (לגבי משתנים מסוימים ובפרט אלה הנבדקים) כך טעות הדגימה בהיסק מהמדגם לאוכלוסייה תהיה קטנה יותר.

את מה דוגמים?
אנו דוגמים את יחידות הניתוח מתוך אוכלוסיית המחקר.
יחידות הניתוח יכולות להיות, למשל, אנשים, כתבות, פרסומות, תוכניות, טלוויזיה, ארגונים, בתי ספר.

הגדרת האוכלוסייה:
לפני כל דגימה יש להגדיר את אוכלוסיית המחקר.
הגדרת האוכלוסייה צריכה להיות מנוסחת במונחים אופרציונליים, כך שיהיה ברור מי נכלל בה ומי לא.
הגדרת האוכלוסייה תכלול את יחידת הניתוח, היקף האוכלוסייה ואת הזמן ו/או המקום שבו נערך המחקר.

קביעת גודל המדגם:

1. הומוגניות / הטרוגניות האוכלוסייה – כאשר אוכלוסייה הומוגנית במשתנה הנחקר, אין חשיבות לגודל המדגם או לצורת הדגימה. מידת ההטרוגניות (שונות) קובעת באיזו מידה ניתן להסיק מן התוצאה במדגם על המצב באוכלוסייה.

2. טעות הדגימה שהחוקר קבע – קיים קשר הפוך בין טעות הדגימה לגודל המדגם. ככל שהחוקר יקבע טעות דגימה קטנה יותר כך יצטרך לדגום מדגם גדול יותר (אם ידגום את כלל האוכלוסייה לא תהיה טעות דגימה כלל).

שיטות דגימה:

שיטות הדגימה נחלקות לשני סוגים:

1. הסתברותית.
2. לא הסתברותית.

דגימה הסתברותית:

מתרחשת אם מתקיימים 3 תנאים:

א. לכל הפריטים באוכלוסייה יש הסתברות ידועה להיבחר למדגם.
ב. לשום פריט אין סיכוי ודאי להיבחר.
ג. שום פריט אינו יוצא מכלל אפשרות להיבחר.

שיטות דגימה הסתברותיות מאפשרות הסקה הסתברותית מהמדגם לאוכלוסייה, תוך חישוב טעות הדגימה.

על מנת לבצע דגימה הסתברותית חייבת להיות ברשות החוקר מסגרת דגימה.

מסגרת דגימה כוללת את כל הפריטים שמתוכם נלקח המדגם הלכה למעשה. היא תהיה קטנה מהאוכלוסייה הנחקרת, ובכל אופן לא יכולה להיות גדולה ממנה. ישנו היסק סטטיסטי – מהמדגם למסגרת הדגימה והיסק אינטואיטיבי – ממסגרת הדגימה לאוכלוסייה הנחקרת. דיוקם של ההיסקים תלויים במידת הייצוג של המדגם את מסגרת הדגימה ושל מסגרת הדגימה את האוכלוסייה.

שיטות דגימה הסתברותיות

דגימה הסתברותית היא דגימה בה לכל פריט יש הסתברות ידועה להיכלל במדגם. ההסתברויות לכל פריט לא חייבות להיות שוות, אך עליהן להיות ידועות, קטנות מאחד וגדולות מ-0, כדי שההסתברות השונה תוכל להיכלל בחישוב.

א. דגימה מקרית פשוטה.
ב. דגימה שיטתית.
ג. דגימת שכבות.
ד. דגימת אשכולות.

דגימה מקרית פשוטה – היא דגימה בה לכל פריט/אובייקט מתוך האוכלוסייה קיים סיכוי שווה להיבחר למדגם. שיטה זו נעזרת בלוחות מספרים אקראיים או בתוכנת מחשב מתאימה הפולטת מספרים באופן מקרי. שיטה זו מקובלת בעיקר באוכלוסיות קטנות יחסית, כאשר מסגרת הדגימה ממוספרת וקל לאתר את

הפריטים השונים לפי מספריהם.

דגימה שיטתית – משמשת בעיקר במצבים בהם מסגרת הדגימה או המדגם גדולים מאוד או כאשר מסגרת הדגימה איננה ממוספרת (כמו למשל, מסגרת דגימה המבוססת על ספר טלפונים) ולכן לא ניתן להשתמש בלוח המספרים המקריים. צריך רשימה מסודרת של מסגרת הדגימה. רק הפריט הראשון נדגם באופן מקרי. שאר הפריטים נדגמים בצורה שיטתית לפי יחס הדגימה – היחס בין גודל מסגרת הדגימה לגודל המדגם הרצוי.

הבעיות בשיטת הדגימה השיטתית:

1. כאשר מסגרת הדגימה זוהי רשימה ממוינת / מסודרת לפי סדר מסוים שרלוונטי לתכונה הנמדדת במחקר.
2. כאשר הרשימה שהתקבלה היא רשימה מחזורית ויחס הדגימה חופף ליחס המחזוריות או לכפולות שלו. במקרה זה יידגם כל פעם אותו פריט וכל שאר הפריטים בתוך המחזור יפוספסו.

דגימת שכבות – כאשר קיימת ידיעה באשר לקיומם של משתני רקע (כמו למשל, גיל, מוצא, מגדר, רמת השכלה וכו') שהם רלוונטיים לתכונה הנמדדת במחקר, אזי רצוי קודם כל לחלק את האוכ' לשכבות לפי כל משתנה רקע רלוונטי בנפרד.

כל שכבה היא הומוגנית בתוכה בתכונה הנחקרת אך הטרוגנית ביחס לשכבות אחרות.
חסרון השיטה – לא ניתן להשתמש בה לשיטות מורכבות לניתוח סטטיסטי.

כמה נבדקים צריך לדגום מכל שכבה?
כדי לדעת כמה נבדקים צריך לדגום מכל שכבה, החוקר צריך לדעת את ההתפלגות (באחוזים) של ערכי המשתנה באוכלוסייה (למשל, התפלגות הנשים והגברים באוכ' הנחקרת).
כלומר, צריך לדעת האם לבצע בדיקת שכבות פרופורציונלית או לא פרופורציונלית.

דגימת שכבות פרופורציונלית – אם קיימים נתונים לגבי ההתפלגות (באחוזים) של ערכי המשתנה באוכלוסייה – החוקר יחשב את מספר הנבדקים שהוא צריך לדגום מכל שכבה (על פי האחוזים וגודל המדגם הרצוי). הנבדקים ידגמו מכל שכבה בדגימה הסתברותית פשוטה.

דגימת שכבות לא פרופורציונלית – כאשר לא קיימים נתונים לגבי ההתפלגות של ערכי המשתנה באוכ' אז פשוט נדגום מכל שכבה מספר נבדקים שווה.
ניתן גם להגדיל / להקטין באופן מלאכותי דגימה של שכבות מסוימות מתוך האוכלוסייה לפי צרכי המחקר.

דגימת אשכולות – משמשת כאשר מסגרת הדגימה גדולה במיוחד ומפוזרת גיאוגרפית או במקרים שאין רשימה אחת מסודרת. כמו למשל במחקרים על בתי ספר, בתי חולים או ארגונים בעלי סניפים מרושתים באזורים שונים.
בשיטה זו מחלקים את האוכלוסייה לתתי-קבוצות (אשכולות) – דוגמים מקרית מספר אשכולות וכוללים במדגם את כל הפריטים המשתייכים לאשכולות שנדגמו. ישנה מקריות בבחירת האשכולות עצמם.

ישנה גם דגימת אשכולות רב שלבית – הליך דגימה הסתברותי בו מבצעים מספר שלבים של דגימת אשכולות.
הבעיה בדגימת אשכולות בכלל וקל וחומר בדגימת אשכולות רב שלבית היא הגדלת טעות הדגימה בשל הדגימה הרב שלבית-של אשכול אח"כ דגימה בתוך האשכול ואח"כ דגימה של פרטים.
היתרון – חיסכון במשאבים וזמינות של רשימות דגימה מצומצמות יותר, נוחות יותר לשימוש וקלות יותר להשגה.

ההבדלים בין דגימת שכבות לדגימת אשכולות:

לשתי השיטות שני שלבים: בשלב הראשון – מחלקים את האוכלוסייה לקבוצות (שכבות או אשכולות) ובשלב השני – דוגמים מכל קבוצה (או את הקבוצות עצמן – במקרה של דגימת אשכולות).

לשתי השיטות מרכיב של בחירה מקרית. בדגימת שכבות הבחירה המקרית היא מתוך השכבות ואילו בדגימת אשכולות הבחירה המקרית היא בבחירת האשכולות עצמן.

	דגימה מקרית	טעות הדגימה	כיצד נקטין את טעות הדגימה
שכבות	מתוך כל שכבה	שונות בתוך השכבות	הומוגניות בתוך כל שכבה הטרוגניות בין השכבות
אשכולות	של אשכולות	שונות בין אשכולות	הומוגניות בין האשכולות הטרוגניות בתוך כל אשכול

<u>המשמעות</u>: כאשר חוקרים משתנה שיש סבירות גבוהה לשונות גדולה יותר <u>בתוך</u> כל אחד מהערכים של המשתנה מאשר <u>בין</u> ערכי המשתנה, נבחר בדגימת אשכולות. לעומת זאת כאשר חוקרים משתנה שקיימת סבירות גבוהה לשונות גדולה יותר <u>בין</u> הערכים השונים מאשר <u>בתוך</u> כל אחד מהערכים, נבחר בדגימת שכבות.

שיטות דגימה רב שלביות

ניתן לשלב את שיטות הדגימה השונות.

למשל:
לצורך סקר על איכות חיים בתל אביב חוקר מחלק שכונות לשכבות על-בסיס סוציו-אקונומי, ואז בכל שכבה מתייחס לשכונות כאשכולות שמתוכם הוא דוגם בצורה מקרית או שיטתית את האנשים שישתתפו בסקר.

דגימה לא הסתברותית:

לפחות אחד מ-3 התנאים הנ"ל אינו מתקיים (1> P ידועה>0).
או שהחוקר מבחין בבעיות העלולות להתעורר בכל אחד משלבי איסוף הנתונים:

אין מסגרת דגימה.

קושי באיתור האנשים (אוכלוסיות סגורות, פעילות לא חוקית).

קושי בשיתוף הפעולה של הנחקרים (נשים מוכות, למשל).

דגימה לא הסתברותית איננה תקפה במחקר מדעי שכן לא ניתן להסיק הסתברותית מהמדגם לאוכלוסייה (לא ניתן לחשב את טעות הדגימה).
אולם ישנם מקרים בהם דגימה הסתברותית אינה אפשרית ולכן קיימות גם שיטות דגימה לא הסתברותיות.

שיטות דגימה לא הסתברותיות:

מדגם נוחות – בחירת הנחקרים הקלים ביותר להשגה על ידי עורך המחקר.
מדגם כזה בד"כ לא מייצג ומקובל במחקרי גישוש.

מדגם כדור השלג – המדגם נבחר בעזרת קבוצה קטנה של מודיעים שמהם מקבלים שמות וכתובות של אנשים נוספים השייכים לקבוצה הנחקרת. משמש כאשר יש קושי להגיע לקבוצה הנחקרת.

מדגם מתנדבים – המדגם מורכב ממתנדבים למחקר. משמש כאשר יש בעיות בנוגע לשיתוף פעולה של המשתתפים (כאשר המחקר מצריך זמן רב למשל).

שיטת המכסה – החוקר בוחר את מי יראיין כשהמגבלה היחידה המחייבת אותו היא מספר המרואיינים מאותה אוכלוסייה. משמש כאשר אין רישום מרכזי של האוכלוסייה כולה או כדי לחסוך בזמן ובמשאבים.

תרגול:

1) חוקר רצה לבחון את השימוש בשירותי בנק אוטומטיים לעומת פנייה לפקיד באשנב, בקרב בנים ובנות ובקרב קבוצות גיל שונות.
הנח כי רשימת מסגרת הדגימה שבידיו (לקוחות רשומים של בנק גדול בארץ) כוללת 10,000 איש, המתפלגים מבחינת קבוצות הגיל כדלהלן:
עד 18: 20%.
18-65: 40%.
מעל 65: 40%.
בנוסף, אחוז הנשים במסגרת הדגימה: 40%. כמו כן אין קשר בין קבוצת הגיל למין הנבדק. החוקר מעוניין לדגום 200 איש מתוך מסגרת הדגימה הנתונה.

א. החוקר מאמין כי הן גיל והן מגדר הם גורמים מרכזיים בקביעת השימוש בשירותי בנק אוטומטיים.
כיצד ידגום את הנבדקים בהתחשב במידע זה?
נמק תוך הצגת הרכב המדגם המתאים.

ב. החוקר מאמין כי רק הגיל הוא גורם מרכזי בקביעת השימוש בשירותי בנק אוטומטיים. כיצד יבצע את הדגימה בהתחשבות במידע זה?
נמק תוך הצגת הרכב המדגם המתאים.

ג. מה מקור טעות הדגימה בכל אחת מהדגימות בסעיפים א' ו-ב'.

2) חוקר רוצה ללמוד מהם חמשת משחקי המחשב האהודים ביותר על ידי ילדים ישראלים מכיתות א', באמצעות שאלון. הוא הצליח לקבל ממשרד החינוך רשימה עם שמותיהם של 25,000 תלמידי כיתות א' מ-450 בתי ספר יסודיים ברחבי הארץ.

א. באיזו שיטת דגימה כדאי לו לבחור? נמק והסבר.

ב. מה מקור טעות הדגימה בשיטה שנבחרה?

3) לגבי כל אחד מהתיאורים הבאים רשמו מהי שיטת הדגימה המתאימה ביותר, מהן יתרונותיה והאם צריך ידע מוקדם לשיטת דגימה זו:

א. חוקר רוצה לבדוק באילו סוגי עבודות נוטים להשתלב נערים ונערות שנשרו ממסגרות לימודיות. הוא מחליט לבחור מדגם של 400 מבין 6,000 בני הנוער שנשרו בשנה הקודמת ממסגרות לימודיות. החוקר מאמין שהגיל בו נושרים מביה"ס הוא הגורם המרכזי בבחירת סוג העבודה. ולכן מעוניין להבטיח שבמדגם יהיה ייצוג הולם לבני הגילאים השונים. החוקר יודע את אחוז כל גיל ביחס לאוכלוסייה הכללית.

ב. חוקר רוצה לבצע סקר על הרגלי הקריאה של האוכלוסייה. ברשות החוקר אין רישום של כל מנויי הספריות העירוניות בארץ, אך יש רישום של כל הספריות.

ג. במחקר על עיירות פיתוח החוקרים נדרשו לכלול 70% תושבי עיירות פיתוח, ו-30% תושבי ערים אחרות.

ד. חוקר מעוניין לבחון מניעים של אנשים הנוהגים לגנוב מחנויות.

4) חוקר רצה לבדוק את שביעות הרצון של אנשים המשתמשים בתחבורה הציבורית. הוא הגיע לתחנת אוטובוס ותחנת רכבת, דגם קרון אחד ברכבת ואוטובוס אחד בתחנת אוטובוסים. בחלק הקדמי בחר מקרית 3 אנשים, ובחלק האחורי בחר מקרית 3 אנשים.

א. מהי יחידת הדגימה / ניתוח במחקר זה?

ב. באיזה שיטה/ות דגימה השתמש?

תשובות סופיות:

1) א. דגימת שכבות פרופורציונאלית לפי משתנים: גיל ומגדר.
ב. דגימת שכבות פרופורציונאלית לפי משתנה: גיל.
ג. סעיף א': בתוך קבוצת גיל ומגדר, סעיף ב': בתוך קבוצת גיל.

2) א. דגימת אשכולות.
ב. בין האשכולות.

3) א. שכבות פרופורציונאלית, יתרונות: ייצוג למשתנה רקע רלוונטי, נדרש: כן.
ב. אשכולות, יתרונות: ביצוע דגימה הסתברותית, נדרש: כן.
ג. שכבות לא פרופורציונאלית, יתרונות: ייצוג לשכבת מיעוט, נדרש: מסגרת דגימה.
ד. כדור שלג, יתרונות: חקירת נושא רגיש עם קושי באיתור אוכלוסייה רלוונטית, נדרש: "מודיעים".

4) א. אנשים המשתמשים בתחבורה ציבורית.
ב. דגימת שכבות (סוג תחבורה), דגימת אשכולות (קרון ואוטובוס), דגימת שכבות (מקום ישיבה), מקרית פשוטה (נוסעים).

שיטות מחקר כמותניות

פרק 4 - איכות כלי מדידה

תוכן העניינים

1. איכות כלי מדידה 27

איכות כלי מדידה:

הקדמה:

רקע:

כלי מדידה – הביטוי הכמותי של המשתנה התיאורטי. מאפשר למדוד כמותית את המשתנה התיאורטי.
איכות כלי המדידה – תלויה במידת ההלימה בין המשתנה התיאורטי, אותו הכלי נועד למדוד, למשתנה התצפיתי שהוא התוצאה המתקבלת מכלי המדידה.

איכות כלי המדידה מוערכת על ידי מדדים של: מהימנות ותוקף.

כלי מדידה הוא תמיד שילוב של שיטה אחת + תכונה אחת.

שיטה: הכוונה לאופן איסוף הנתונים (שאלון, ראיון, מבחן, תצפית).
תכונה: המשתנה התיאורטי שרוצים לבדוק (למשל רמת תוקפנות).

על מנת לבדוק את איכות כלי המדידה יש לבדוק את המהימנות והתוקף של הכלי:
מהימנות כלי המדידה – נשאלת השאלה עד כמה כלי המדידה שיטתי, עקבי.
במובן הזה שאם נחזור על המדידה פעם נוספת נקבל את אותן התוצאות.
תוקף כלי המדידה – נשאלת השאלה האם הכלי אכן בודק את מה שהוא אמור לבדוק, כלומר האם המשתנה התצפיתי אכן בודק את המשתנה התיאורטי ואך ורק משתנה זה.

מהימנות:

רקע:

עד כמה ניתן לסמוך על כלי המדידה שהוא מודד ערכו של המשתנה בדייקנות ועד כמה מושפע מטעויות מקריות.

טעות מדידה:

טעות שיטתית – הכלי מזייף בגודל קבוע ואחיד לאורך כל הנבדקים וכל המדידות: $\sigma_e^2 = 0$.

טעות מקרית – גודלה משתנה ממדידה למדידה ומנבדק לנבדק: $\sigma_e^2 \neq 0$.

מהימנות הכלי בודקת את העקביות של הכלי ולכן נפגעת רק מטעות מקרית (ולא נפגעת מטעות קבועה ששונותה = 0).

מקדם המהימנות - מבטא יחס בין השונות האמיתית לשונות הנצפית:

$$r_{tt} = \frac{\sigma_t^2}{\sigma_o^2} = r_{to}^2$$

המדידה מהימנה יותר ככל שהשונות בין האנשים מוסברת יותר על ידי הבדלים אמיתיים ופחות על ידי טעויות מקריות.

$$0 \leq r_{tt} \leq 1$$

מהימנות מושלמת	אין מהימנות
אין גורם טעות כלל	הכלי לא מבטא כלל את הערך האמיתי
שונות נצפית = שונות אמיתית	שונות נצפית = שונות הטעויות

אם למשל $r_{tt} = 0.55$ המשמעות היא כי המדידה משקפת 55% שונות אמיתית ו-45% שונות שמקורה בטעויות מקריות.

כלי נחשב מהימן אם $r_{tt} \geq 0.7$, תלוי באופן שבו המהימנות נמדדה.

שורש המהימנות ($\sqrt{r_{tt}} = r_{to}$) מהווה חסם עליון לכל מתאם בין כלי המדידה לכל כלי מדידה אחר: $r_{tt} \geq r_{xy}$.

אם מהימנותו של הכלי נמוכה זה אומר שהוא לא מתואם עם עצמו אז קל וחומר שלא יוכל להיות מתואם עם כלים אחרים.

המשמעות:

המהימנות משפיעה על יכולת הניבוי של הכלי. ככל שהמהימנות נמוכה כך גם יכולת הניבוי של הכלי תהיה נמוכה ולהיפך.

הגורמים לטעות המקרית של המדידה:

1) זמן המדידה – גורמים נסיבתיים העלולים להשפיע באופן שונה על כל נבדק:
   א. גורמים אקראיים – חיצוניים – קשורים במזל (כגון: מזג האוויר, תנאים פיזיים של המקום, טמפ' החדר, מידת הנוחיות, כמות הרעש, מצב התאורה וכו').
   ב. גורמים אישיים זמניים – קשורים בנבדקים עצמם (כגון: מצב הרוח, רמת ריכוז, מוטיבציה, עייפות, מצב בריאות וכו').
2) תוכן הפריטים – ייצוגיות הדגימה של עולם התוכן – האם יש דגימה טובה של פריטי המבחן כך שייצגו בצורה מהימנה את עולם התוכן של המבחן? האם ניסוחם ברור ומובן באותו אופן לכל הנבדקים?
3) סובייקטיביות הבודק – רלוונטיים למבחנים בהם הציון איננו אובייקטיבי אלא קשור לבודק מסוים. עד כמה הציון חשוף לשיקול דעתו ופרשנותו הייחודית של כל מעריך?

דרכים לבדיקת מהימנות הכלי:

בדיקות שונות הבאות לבדוק את רגישות הכלי למקורות הטעות המקרית.

1. מהימנות כיציבות.
2. מהימנות כאקוויוולנטיות.
3. מהימנות כעקיבות פנימית.
4. מהימנות בין שופטים.

מהימנות כיציבות:

מכונה מהימנות מבחן חוזר (Test-Retest): מקדם המהימנות מבטא את המתאם בין שתי העברות של אותו מבחן בהפרש של זמן: $r_{tt} = r_{t_1 t_2}$.
ההנחה היא שהמבחן מהימן כאשר תוצאותיו יציבות לאורך זמן.
בשיטה זו נבדקת רגישות המבחן לגורמים הקשורים לזמן המדידה – גורמים נסיבתיים – חיצוניים או פנימיים. כדי שהמבחן יחשב מהימן בשיטה זו: $t_{tt} \geq 0.8$.
ככל שהפער בזמן גדל, כך המהימנות שתתקבל תהיה נמוכה יותר.

מהימנות כאקוויוולנטיות:

מכונה מהימנות נוסחים מקבילים (Parallel Forms): מקדם המהימנות מבטא את המתאם בין שני נוסחים מקבילים של אותו המבחן: $r_{tt} = r_{AB}$.
ההנחה היא שהמבחן מהימן אם פריטי התוכן שבו מיוצגים בצורה טובה. מבחן זה רגיש לטעויות מדידה מקריות שנבעו מחוסר ייצוגיות הדגימה (תוכן, אופי וניסוח הפריטים).

קיימים שני סוגים של מהימנות כאקוויוולנטיות:

1. מבחן מקביל בו זמנית – שני הנוסחים מועברים באותו הזמן.
   נצפה ל: $r_{tt} \geq 0.8$.
2. מבחן מקביל לאחר זמן – שני הנוסחים מועברים בהפרש של זמן.
   נצפה ל: $r_{tt} \geq 0.7$.

המדד השני רגיש פרט לייצוגיות הדגימה גם לגורמים נסיבתיים - אקראיים שעלולים היו להשפיע על תוצאותיו (בגלל הפרש הזמנים).
אם נקבל מהימנות גבוהה נדע כי המבחן מתגבר על שתי הבעיות אולם במידה והמהימנות נמוכה לא ניתן לדעת האם הבעיה היא בייצוגיות הדגימה או בגורמים הנסיבתיים.

מהימנות כעקיבות פנימית:

קיימים שני סוגים של מהימנות כעקיבות פנימית:

1. מבחן חצוי.
2. קרונבך אלפא / קודר-ריצ'רדסון.

מבחן חצוי:

בדיקת מהימנות המבוססת על מתאם בין שני חלקים של המבחן. שלבי עבודה:

- חלוקה מקרית של המבחן לשני חלקים.
- חישוב מתאם בין שני החלקים.
- הצבת הערך שהתקבל בנוסחת ספירמן-בראון: $r_{tt} = \frac{2r_{1,2}}{1+r_{1,2}}$.

מהימנות המבחן כולו תהיה גדולה יותר מהמתאם בין שני חלקי המבחן (הנוסחה מבצעת תיקון כלפי מעלה של המתאם בין שני החלקים).

אלפא קרונבך:
מתבסס על ממוצע כל המתאמים הזוגיים האפשריים בין הפריטים (מדד קודר ריצ'רדסון מחושב עבור משתנים בסולמות דיכוטומיים). מדד זה מבטא את ממוצע המתאמים בין כל שני פריטים בשאלון. ככל שהמבחן ארוך יותר כך יהיה מהימן יותר כי דוגמים שוב ושוב את אותו המשתנה.
מדד זה בודק פרט לייצוגיות הדגימה גם את מידת ההומוגניות של המבחן.

מהימנות בין שופטים:
מדד מהימנות זה רלוונטי רק במדידות בהן ערכי המשתנה התצפיתי נקבעות באופן סובייקטיבי על ידי שופט. מקדם המהימנות מבוסס על המתאם או ממוצע המתאמים בין הערכות שופטי בלתי תלויים לביצוע מסוים: $r_{tt} = r_{ab}$.
הערך המצופה: $r_{tt} \geq 0.7$. הטעות נובעת מסובייקטיביות המעריך הכוללת כל דבר שקשור לבודק כמו מצב רוחו, עיוותי תפיסה, אפקט הילה, סטריאוטיפים וכו'.
2 השופטים חייבים להיות בעלי אותה הזיקה לקבוצת הנבדקים.

סוגי מבחנים ומדדי המהימנות המתאימים להם:

- מבחן הומוגני – פריטי המבחן מודדים גורם יחיד – ניתן להשתמש בכל סוגי המהימנויות.
- מבחן הטרוגני – פריטי המבחן מודדים היבטים שונים של תופעה מסוימת – לא ניתן להשתמש במדד אלפא של קרונבך.
- מבחן עוצמה – רמת קושי עולה – מדד קרונבך אלפא לא מתאים כי מצפים למתאמים נמוכים בין שאלות מרמות קושי שונות.
- מבחן מהירות – הגבלת זמן – לא מתאים עקיבות פנימית כי יש מספר שונה של פריטים שאנשים הצליחו לענות עליהם.
- מבחנים אובייקטיביים – בעלי מפתח ברור למתן ציונים שאינו תלוי בשיפוט סובייקטיבי של המעריך – מדד מהימנות בין שופטים איננו רלוונטי.
- סוג התכונה – אם התכונה אינה יציבה לא ניתן להשתמש במהימנות כיציבות ובמהימנות כאקוויוולנטיות מדד ב.

<u>הערה חשובה</u>: על מנת שנזהה שמדובר בבדיקה של מהימנות מסוג כלשהו, יש לזהות שמחושב מתאם בין שתי מדידות שמתבצעות על אותו כלי מדידה, כלומר שחייבים להיות שלושה מרכיבים בכל בדיקת מהימנות:

1. מדובר בקב' נבדקים אחת.
2. מדובר בבדיקת תכונה אחת.
3. באמצעות שיטה אחת.

אם אחד משלושת המרכיבים לא מתקיים, כבר לא מדובר בבדיקת מהימנות.

סיכום – מדדי המהימנות:

סוג מהימנות	מס' העברות	מה משתנה	גורם הטעות נובע מ-	המדד הסטטיסטי	מהימנות מצופה	חסרון
יציבות – מבחן חוזר	2	זמן ההעברה	גורמים נסיבתיים	מתאם בין שתי העברות של מבחן זהה בהפרשי זמן	$r_{tt} \geq 0.8$	1.זיכרון ולמידה. 2.לא מתאים לתכונות לא יציבות.
אקוויוולנטיות – מבחן מקביל בו זמנית	1	תוכן השאלות	חוסר ייצוגיות הדגימה	מתאם בין העברה של שני מבחנים מקבילים	$r_{tt} \geq 0.8$	קושי לבנות שני נוסחים באותה רמת קושי.
מבחן מקביל לאחר זמן	2	זמן העברה + תוכן השאלות	חוסר ייצוגיות הדגימה + גורמים נסיבתיים		$r_{tt} \geq 0.7$	במדד ב' – לא ניתן לדעת מהיכן הטעות.
עקיבות פנימית – α קרונבך וקודר ריצ'רדסון	1	תוכן השאלות	חוסר ייצוגיות הדגימה + הטרוגניות	נוסחת α קרונבך	$r_{tt} \geq 0.85$	לא מתאים למבחנים הטרוגניים.
מבחן חצוי			חוסר ייצוגיות הדגימה בלבד	חישוב מתאם בין שני חצאי מבחן והצבה בנוסחת ספירמן בראון.	$r_{tt} \geq 0.8$	מדויק פחות מ-α קרונבאך.
בין שופטים	1	המעריך	סובייקטיבית המעריך	מתאם בין הערכות השופטים.	$r_{tt} \geq 0.7$	1.השקעה בקריטריונים 2.אימון של בודקים.

תרגול:

1) חוקר בנה שלושה שאלונים דומים (X, Y, Z) לבדיקת בטחון עצמי בעבודה טיפולית. כל אחד מהשאלונים כלל 40 פריטים. שלושת השאלונים הועברו לקבוצה בת 100 נבדקים, גברים ונשים, באפריל 85 (מועד א׳) ושוב במאי 85 (מועד ב׳).
עבור כל אחד מהמצבים המתוארים להלן, ציינו והסבירו מהי הדרך שנבחרה למדידת המהימנות ומהו האחוז בשונות הציונים הנצפים שניתן לייחס למקור הטעות הרלוונטי:

א. התקבל מתאם של 0.9 בין X ל Y שהועברו במועד הראשון.

ב. התקבל מתאם של 0.86 בין X של המועד הראשון לבין Y של המועד השני.

ג. התקבל מתאם של 0.95 בין Y שהועבר במועד הראשון לבין Y שהועבר במועד השני.

ד. שאלון X של מועד א חולק לשניים. נתקבל מתאם של 0.85 בין שני החלקים.

ה. ממוצע המתאמים של כל החלוקות האפשריות של פריטי שאלון Z היא 0.89.

2) יש לסמן נכון או לא נכון לגבי כל אחת מהטענות הבאות ולנמק:

א. כאשר מתקבל מקדם אלפא נמוך ניתן לדעת בוודאות שהמבחן אינו מהימן.

ב. נבדקה מהימנותו של מבחן חדש למדידת אסרטיביות באמצעות שיטת הנוסחים המקבילים בזמנים שונים. נמצא כי מקדם המתאם בין שני הנוסחים הינו 0.87 מכאן משמעות הדבר היא ש-13% משונות הנצפית ניתן לייחס לטעויות שמקורן בגורמים נסיבתיים בלבד.

ג. נמצא כי ממוצע inter-item correlation שהתקבל לשאלון דימוי עצמי הינו 0.82, מכאן ניתן לומר שהשאלון מהימן והומוגני.

3) חוקרת בנתה מבחן לבדיקת כושר מכני. המבחן כלל שאלות שונות כגון: ״בנסיעה ברכב בזמן בלימה לאן יטולטל קודם גוף הנוסעים, קדימה או אחורה?״. השאלון כלל 25 פריטים, לכל פריט היו שתי תשובות אפשריות. אחת נכונה ואחת לא נכונה. המבחן הועבר ל 50 נקידים. מדד קודר ריצ׳רדסון שהתקבל הינו 0.92.
על מנת לבחון את מהימנות מבחן 1 בנתה החוקרת נוסח דומה (להלן מבחן 2) והעבירה אותו לאותם אנשים לאחר שבוע. היא בחנה את הקשר בין הציונים שנתקבלו ממבחן 1 לבין הציונים שנתקבלו ממבחן 2.

א. איזו בדיקת מהימנות ערכה החוקרת?

ב. המתאם שהתקבל היה 0.7. האם החוקרת צריכה להיות מופתעת שלא קיבלה ערך זהה למדד קודר ריצ׳רדסון?

ג. עוזר המחקר הציע לחוקרת להעביר את מבחן 1 לקבוצה אחרת במועד אחר, בנוסף להעברה הראשונה של מבחן 1. הוא בדק מהימנות מסוג:
   i. יציבות.
   ii. אקוויוולנטיות.
   iii. שילוב של יציבות ועקיבות.
   iv. לא ניתן לדעת.
   v. זו לא בדיקת מהימנות.

4) נבחנה מהימנות מבחן משכל חדשני. לשם כך נאספו נתוני הביצוע של נבדקים מחוננים במבחן בנוסח יחיד של המבחן, לצורך חישוב מהימנות המבחן החצוי. מהימנות המבחן נמצאה נמוכה. הסבר אפשרי לכך הינו:

א. מדובר במבחן הטרוגני מבחינת תחומי התוכן הנחקרים בו.

ב. מדובר במבחן הומוגני מבחינת תחומי התוכן הנחקרים בו.

ג. מדובר במדגם נבדקים הומוגני מבחינת התכונה הנחקרת.

ד. תשובות ב' ו-ג' נכונות.

5) שני נוסחים מקבילים של מבחן לבדיקת מנת משכל הועברו למדגם של 100 נבדקים. לאחר שבוע הועבר הנוסח השני של המבחן שנית. להלן תוצאות חישובי המהימנות:
התקבל מתאם של 0.90 בין שני הנוסחים שהועברו באותו הזמן.
התקבל מתאם של 0.85 בין נוסח א' לבין נוסח ב' לאחר שבוע.
התקבל ערך קרונבך של אלפא 0.88 עבור נוסח א'.
מה אחוז השונות בציונים נצפים שאפשר לייחס לטעויות שמקורן:

א. בייצוגיות דגימת התוכן בלבד.

ב. בגורמים נסיבתיים/אקראיים בלבד.

ג. בייצוגיות דגימת התוכן ובהטרוגניות המבחן.

6) חוקר בנה מבחן לרצייה חברתית והעביר אותו פעמיים בהפרש של חודשיים ולאחר מכן חישב את המתאם בין שתי ההעברות.

א. איזה סוג מהימנות נבדק?

ב. מה בודקים בסוג מהימנות זה?

ג. המתאם שהתקבל היה 0.74, מה ניתן לומר על מהימנות המבחן?

ד. החוקר ביקש לבדוק האם המבחן מושפע מבחירת הפריטים, ניסוחם וסוג הפריטים. באילו סוגי מהימנות הוא יכול להשתמש?

ה. במה עדיף להשתמש?

ו. לתכונת חרדה יש מספר ממדים, איזה סוג מהימנות יחשב החוקר?

ז. אם בכל זאת יחושב סוג המהימנות השני, מה צפוי להיות ערכו?

ח. מהימנות מבחן חצוי=0.82. החוקר טען שהתגבר על גורמי הטעות. האם צדק?

7) כשמשקל א' מראה שהנשקל שוקל 60 ק"ג ומשקל ב' מראה שהוא שוקל 58 ק"ג אפשר לומר:

א. מהימנותו של משקל א נמוכה.

ב. מהימנותו של משקל ב נמוכה.

ג. מהימנותם של שני המשקלים נמוכה.

ד. מהימנותם של שני המשקלים גבוהה.

ה. לא ניתן להסיק מהבדיקה שבוצעה על מהימנות המשקלים.

8) חוקר בנה 3 שאלונים שבודקים דימוי עצמי:
C=היבט חברתי, S=היבט רגשי, G=היבט גופני.
כל אחד מהשאלונים כלל 40 פריטים שהועברו לקבוצה בת 100 נבדקים, גברים ונשים באפריל 2015 (מועד 1) ושוב במאי 2015 (מועד 2).
עבור כל אחד מהמצבים המתוארים להלן, ציין והסבר מהי הדרך שנבחרה לבדיקת המהימנות. אם מדובר בעקיבות פנימית ציין מהי הטכניקה ומהו המקדם שחושב.
ציין מהו חלקה של השונות האמיתית וממה היא משוחררת.

א. $r_{G_1,G_2}=0.8$.

ב. שאלון S של מועד 1 חולק לשניים. נתקבל מתאם של 0.81 בין שני החלקים.

ג. ממוצע המתאמים של כל החלוקות האפשריות של פריטי שאלון S היה 0.89.

ד. עוזר מחקר רצה להעלות את הערך שהתקבל בסעיף הקודם. לצורך זה הוא הוסיף קבוע חיובי לכל ציוני הנבדקים בשאלון S. האם הוספת קבוע זה תשפיע על ערך מקדם המהימנות? הסבר.

ה. מבחן C שהועבר במועד 2 ניתן לקידוד של שני מעריכים בלתי תלויים. המתאם שחושב בין שני המעריכים היה שווה ל-1.

תשובות סופיות:

1) א. אקוויוולנטיות, 10%. ב. אקוויוולנטיות, 14%. ג. מבחן חוזר, 5%.
ד. עקיבות פנימית, 15%. ה. עקיבות פנימית, 11%.

2) א. לא נכון. ב. לא נכון. ג. נכון.

3) א. אקוויוולנטית. ב. לא. ג. v.

4) ג'.

5) א. 10%. ב. 5%. ג. 12%.

6) א. יציבות, מבחן חוזר. ב. רגישות המבחן לטעויות שמקורן בגורמים נסיבתיים.
ג. נמוכה. ד. אקוויוולנטיות או עקיבות פנימית. ה. עקיבות פנימית.
ו. עקיבות פנימית. ז. נמוך. ח. לא.

7) ה.

8) א. יציבות, 80%. ב. עקיבות פנימית-מבחן חצוי, 90%.
ג. עקיבות פנימית, 89%. ד. לא. ה. בין שופטים.

תוקף:

רקע:

תוקף מתאר את המידה שבה כלי המדידה מודד את התכונה שהוא אמור למדוד ולא תכונות אחרות.
מאחר ולא נתון בידינו המשתנה התיאורטי, אנו נבדוק את מידת ההתאמה של הכלי שלנו למשתנה התיאורטי בדרכים עקיפות.

לתוקף שלוש הגדרות אופרטיביות:
תוקף תלוי קריטריון (תוקף ניבוי) – המידה שתוצאות המבחן מאפשרות חיזוי מדויק של מדידה אחרת (קריטריון).
תוקף תוכן – המידה בה פריטי המבחן בודקים את עולם התוכן של המשתנה.
תוקף מבנה – מידת ההתאמה בין המשתנה הנמדד (התצפיתי) למשתנה התיאורטי (התכונה).

תוקף תלוי קריטריון:

מתייחס למצב בו החוקר מעוניין לפתח כלי שניתן יהיה לנבא באמצעותו משתנה מסוים הקרוי קריטריון.
למשל, הפסיכומטרי הוא כלי שנועד לנבא הצלחה בלימודים -הקריטריון.
הכלי (מבחן פסיכומטרי) יחשב בעל תוקף תלוי קריטריון גבוה אם מידת ההתאמה בינו לבין הקריטריון (ציונים בתואר הראשון) תהיה גבוהה.

בדיקת תוקף תלוי קריטריון:

בדיקת התוקף תלויה באופי המשתנים (איכותיים או כמותיים):

1. חישוב מתאם בין מנבא לקריטריון - r_{xy} - מקדם תוקף תלוי קריטריון (כלי וקריטריון כמותיים).
2. טבלת ציפיות (קריטריון איכותי).
3. טבלת שכיחויות (כלי וקריטריון איכותיים).

שיטות תיקוף תלויות קריטריון:

ישנן 3 שיטות תיקוף תלויות קריטריון. שיטות אלו נבדלות בזמני מדידת הציון בכלי (מנבא) ביחס לציון בקריטריון (מנובא):

1. תוקף לאחר מעשה – קריטריון נמדד לפני הכלי.
2. תוקף בו זמנית – הקריטריון והכלי נמדדים בו זמנית.
3. תוקף הניבוי – הקריטריון נמדד אחרי הכלי.

תרגול:

1) מהו סוג תוקף תלוי קריטריון במקרים הבאים:

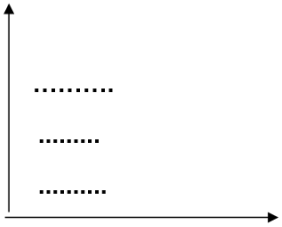
א. תלמידים לקראת קבלה בבית ספר למשפטים נדרשים לעבור מבחן באנגלית. חושב מתאם בין ציוני המבחן לבין ממוצע הציוני שנה א' של סטודנטים שהתקבלו באותה שנה ללימודים.

ב. קבוצה של חולי לב וקבוצה של אנשים בריאים התבקשו למלא שאלון לבדיקת סיכון ללקות במחלת לב. חושב הקשר בין ציוני השאלון לבין שתי הקבוצות.

ג. עובדים בחברת הייטק התבקשו לענות על שאלון לבדיקת יכולת מנהיגות והממונים עליהם נתבקשו למלא שאלוני הערכה על ביצועי העובדים. מתאם חושב בין ציוני השאלונים.

ד. חוקר בנה מבחן שמטרתו לאתר את המתאימים להכשרה לתפקיד דיפלומטי במשרד החוץ. קבוצה של 90 מועמדים נבחנה במבחן. כל המועמדים התקבלו לקורס בלי קשר להישגיהם במבחן. בתום הקורס נבחן הקשר בין תוצאות במבחן (מתאים/לא מתאים) לבין ציון סוף הקורס (עבר/נכשל).

2) הקשר בין ציוני מבחן להכשרה לתפקיד דיפלומטי לבין ציון סוף קורס מוצג בתרשים:



האם יש למבחן תוקף ניבוי?

תשובות סופיות:

1) א. בו זמני. ב. לאחר מעשה. ג. בו זמני. ד. ניבוי.

2) לא.

רקע:

תוקף תוכן:

מבטא את המידה שבה פריטי המבחן בודקים את עולם התוכן של המשתנה הנמדד. האם פריטי המבחן רלוונטיים לעולם התוכן הנמדד והאם מייצגים את כל היבטיו? כדי להבטיח תוקף תוכן אנחנו צריכים להקפיד על ניתוח טוב של עולם התוכן בו אנו מעוניינים, צריכים לבחור אוסף מייצג של פריטים מעולם התוכן הזה ובנייה טובה (סדר השאלות, למשל) של מכשיר המדידה (השאלון / מבחן).

תהליך התיקוף של תוקף תוכן:
תהליך התיקוף הוא התרשמותי-שיפוטי ומתרחש בשלב הבנייה של הכלי ואמורים להשתתף בו מומחים ויועצים בתחום.
תהליך התיקוף יכלול את השלבים הבאים:

1. ניתוח שיטתי של עולם התוכן בו מעוניינים.
2. בניית אוסף מייצג של הפריטים.
3. שיפוץ הפריטים בתהליך דיון והסכמה.

תוקף מבנה:

מתייחס למידה שבה המבחן מודד את התכונה / המבנה התיאורטי. כלומר, למידת ההתאמה בין המשתנה התצפיתי למשתנה התיאורטי.

למשל:

משתנה תיאורטי: רמת תוקפנות

↓

הגדרה נומינאלית: התנהגות אלימה של אדם

↓

הגדרה אופרציונלית: מדידת כמות המילים השליליות בשיחה

↓

משתנה תצפיתי: מספר המילים השליליות בשיחה.

תהליך התיקוף של תוקף מבנה:
תהליך ארוך יותר מהקודמים וכולל את כל סוגי הבדיקות של התוקף והמהימנות שהוזכרו.

הדרכים לבחינת תוקף מבנה:

1. קורלציות עם משתנים קשורים: בדיקת מתאמים של אותו מושג עם מושגים דומים.
2. הבדלים בין קבוצות: החוקר יבחן את המידה שבה מתגלים הבדלים בין קבוצות שההבדל הבולט ביניהן הוא בתכונה הנמדדת ולהשוות את הציונים שלהן במכשיר המדידה.
3. תוקף הניבוי כחלק מתוקף מבנה.
4. בדיקת מהימנות כחלק מתוקף מבנה.

דרכים אלו לתוקף מבנה מתבססים על שיטתם של **קרונבך ומיל (דרכים לפיתוח תיאוריה).**

5. תוקף מתכנס: נמדוד את אותו המושג בדרכים שונות (שיטות שונות). אם המתאמים המתקבלים הם דומים, נסיק על תוקף מבנה.
6. תוקף מבחין: בדיקת מתאמים עם מושגים שאינם קשורים.
אם המתאמים נמוכים, נסיק שיש אפשרות שמדובר במושג חדש.

שתי הדרכים הללו לתוקף מבנה מתבססים על שיטתם של **קמבל ופיסק (מטריצת תכונות שיטות).**

מטריצת MMM (מטריצת שיטות תכונות):

כלי מדידה = תכונה + שיטה.
הנתונים במטריצה מתייחסים למקדמי המהימנות והתוקף (מבחין ומתכנס) של שני כלים שונים:

		שיטה 1 (שאלון)		שיטה 2 (ריאיון)	
		תכונה 1 (חרדה)	**תכונה 2 (פחד)**	**תכונה 1 (חרדה)**	**תכונה 2 (פחד)**
שיטה1 (שאלון)	**תכונה 1 (חרדה)**	מהימנות			
	תכונה 2 (פחד)	מבחין א	מהימנות		
שיטה 2 (ריאיון)	**תכונה 1 (חרדה)**	מתכנס	מבחין ב	מהימנות	
	תכונה 2 (פחד)	מבחין ב	מתכנס	מבחין א	מהימנות

כללים:

1. ערכי המהימנות – אותה תכונה הנבדקת באותה שיטה – צריכים להיות הערכים הגבוהים בטבלה.
2. ערכי התוקף המתכנס – אותה תכונה הנבדקת בשיטות שונות – צריכים להיות גבוהים.
3. ערכי תוקף מבחין א' – תכונות שונות הנבדקות באותה השיטה – נמוכים מהמתכנס.
4. ערכי תוקף מבחין ב' – תכונות שונות הנבדקות בשיטות שונות – הנמוכים ביותר בטבלה.

ניתוח המטריצה צריך להתייחס בראש ובראשונה למהימנות. אם שיטה איננה מהימנה אז כל מדדי התוקף שייבחנו עם אותה השיטה יהיו מוטלים בספק.

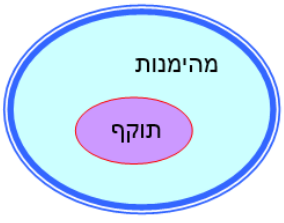
דרך להבחין בין בדיקת תוקף לבדיקת מהימנות:
מהימנות: אותה שיטה ואותה תכונה.
תוקף: שיטות שונות ו/או תכונות שונות.

סיכום תוקף – שיטות ותכונות:

סוג התוקף	תכונות	שיטות	קבוצות	
תוקף תלוי קריטריון	1	2	1	מנבא וקריטריון
חקר הבדלים בין קבוצות	1	2	2	מכשיר המדידה והשתייכות קבוצתית
חקר מתאמים	2 שונות	1 או 2	1	
תוקף מתכנס	1	2	1	
תוקף מבחין	2 דומות	1 או 2	1	

קשר בין מהימנות לתוקף:
אם במהימנות אנו בודקים את רגישות הכלי לטעויות מקריות, בתוקף אנו בודקים את רגישות הכלי לטעויות קבועות ושיטתיות.
לדוגמא: משקל שמראה תמיד קילו נוסף – מהימן אך לא תקף.

1. **מהימנות היא תנאי הכרחי אך לא מספיק לתוקף.**
כדי שמבחן יהיה תקף הוא חייב להיות מהימן אולם מבחן מהימן הוא לא בהכרח תקף.



2. **מהימנות מהווה חסם עליון לתוקף** – שורש המהימנות מהווה חסם עליון לקשר שבין כלי המדידה לכל כלי מדידה אחר. ערך התוקף יגיע לכל היותר לערך שורש המהימנות שחושבה.
3. **הקשר בין מהימנות לתוקף תוכן** – אם העקיבות הפנימית גבוהה יש אישוש לתוקף תוכן (אלא אם כן המבחן הטרוגני).
4. **הקשר בין מהימנות ותוקף מבנה** – התיאוריה מכתיבה את מידת העקיבות הפנימית ומידת היציבות של התכונה לאורך זמן.

תרגול:

מטריצת MMM:

1) לפניך טבלה המסכמת מתאמים בין ציונים שהתקבלו בארבעה כלי מדידה שונים, וכן נתונים לגבי מהימנותם:

		שאלון		ראיון	
		אסרטיביות	**מנהיגות**	**אסרטיביות**	**מנהיגות**
שאלון	**אסרטיביות**	0.85			
	מנהיגות	0.46	0.86		
ראיון	**אסרטיביות**	0.62	0.28	0.84	
	מנהיגות	0.28	0.40	0.65	0.89

התוקף המתכנס, התוקף המבחין החד-שיטתי והתוקף המבחין הרב-שיטתי של ראיון אסרטיביות הם (לפי סדר זה):

א. 0.62 ; 0.46 ; 0.28.
ב. 0.62 ; 0.65 ; 0.28.
ג. 0.85 ; 0.62 ; 0.21.
ד. 0.46 ; 0.62 ; 0.21.

2) פותח ראיון פסיכיאטרי חצי-מובנה לאבחון דיכאון. במחקר המשך, נבחנה מהימנותו, וכן קשרים בין ההערכות המתקבלות באמצעותו לבין כלי מדידה אחרים. הממצאים מסוכמים במטריצה שלפניך (הערכים בטבלה הם תוצאות חישוב המתאמים הרלוונטיים):

		דיכאון		סכיזופרניה	
		סולם MMPI	ראיון פסיכיאטרי	סולם MMPI	ראיון פסיכיאטרי
דיכאון	סולם MMPI	0.86			
	ראיון פסיכיאטרי	0.25	0.87		
סכיזופרניה	סולם MMPI	0.54	0.18	0.84	
	ראיון פסיכיאטרי	0.36	0.23	0.26	0.88

מה ניתן ללמוד מן הנתונים שהתקבלו במחקר (המפורטים בטבלה דלעיל) על תוקפו של כלי המדידה שפותח? התייחס לכלל הנתונים הרלוונטיים. נסה להציע הסבר לממצאים.

תוקף:

3) במטרה לברור מועמדים מתאימים לביה"ס לפקידות, הועבר למועמדים מבחן מטלות פקידותיות X1 שבוחן את מידת הדיוק בהעתקת מסמך. כל המועמדים התקבלו לביה"ס. לאחר סיום הלימודים ניתן לכל נבדק ציון סופי Y שביטא את מידת ההצלחה בלימודים. בדיקת המתאם בין X1 ל-Y היא בדיקת:

א. תוקף מתכנס של X1.

ב. תוקף ניבוי של X1.

ג. תוקף מבחין של X1.

ד. תוקף תוכן של Y.

4) כשפותחו מבחני האינטליגנציה, על מנת לתקפם, בחנו ילדים במבחן אינטליגנציה וביקשו גם מהמורה שתיתן לכל ילד ציון אינטליגנציה על פי היכרותה עימו. החוקרים חישבו מתאם בין שתי מערכות ציונים אלו. במקרה זה ניתן לומר כי:

א. נבדק תוקף התוכן של מבחן האינטליגנציה כחלק מתוקף המבנה של המבחן.

ב. כחלק מבדיקת תוקף המבנה של מבחן האינטליגנציה נבדק תוקף הניבוי של המבחן, הקריטריון: הערכת המורה.

ג. נבדק תוקף הניבוי של הערכת המורה. הקריטריון: ציון במבחן.

ד. נבדק התוקף המבחין של מבחן האינטליגנציה כחלק מבדיקת תוקף המבנה של המבחן.

ה. אף תשובה אינה נכונה.

5) חוקר בנה שאלון למדידת כשרון מוזיקלי. כדי לבדוק את תוקפו המתכנס נתן לכל נבדק ציון בכישרון מוזיקלי על ידי מומחה שהאזין לקטע שהנבדק ניגן. המהימנות שחישב החוקר לשאלון הייתה קרובה ל-0. מכאן ניתן להסיק:

א. התוקף המתכנס (מתאם בין השאלון לציון על הקטע המוזיקלי) יהיה גם הוא קרוב לאפס.

ב. התוקף המתכנס יהיה קרוב ל-1.

ג. על סמך הנתונים אי אפשר לדעת מה יהיה ערכו של התוקף המתכנס.

ד. מקדם התוקף המתכנס יכול להיות כל מספר.

6) פסיכולוג בנה שאלון למדידת גבריות, והעביר אותו בנפרד לקבוצת גברים ולקבוצת נשים, לצורך חישוב ההבדלים בין הקבוצות. השיטה בה השתמש הפסיכולוג לתיקוף השאלון מקבילה להשוואת הביצוע:

א. של גברים ושל נשים במבחן אינטליגנציה.

ב. של מדריכים ושל חניכים בקורס טייס, במבחן אינטליגנציה.

ג. של נשים ושל גברים במבחן טייס.

ד. של מדריכים ושל חניכים בקורס טייס, במבחן טייס.

7) אילו מבחן הנהיגה בארץ היה כולל שאלות במתמטיקה דיפרנציאלית ואינטגראלית, אילו מהטענות הבאות הייתה נכונה:

א. המבחן אינו אמין.

ב. המבחן אינו תקף.

ג. המבחן אינו מהימן.

ד. המבחן אינו מבחן הישג.

8) הקשר שבין מבחן הכניסה הפסיכומטרי לבין הישגי התלמידים בסיום תואר ראשון, מודדת את ה _____ של המבחן הפסיכומטרי :

א. מהימנות.

ב. תקיפות.

ג. מהימנות ותקיפות.

ד. אף אחד מהנ"ל.

9) מרפאה לבריאות הנפש סובלת מבעיה של חוסר בכוח אדם, ולכן היא מעוניינת בסלקציה בפונים לפי הסיכוי שלהם להצליח בטיפול. נמצא מתאם של 0.2 בין מבחן שבדק את מידת ההתאמה של נבדק לטיפול לבין נשירת הנבדקים מהטיפול. סמן/י את המשפט הנכון ביותר :

א. למבחן יש תוקף ניבוי.

ב. למבחן אין תוקף ניבוי.

ג. למבחן יש מהימנות גבוהה.

ד. תוקף הניבוי גבוה אך לא מהימנותו של המבחן.

ה. אף אחד מהנ"ל.

10) חוקר רצה להחליט האם להשתמש בשאלון רגישות (X1) ובשאלון יכולת אמפאטית (X2) לקבלת מועמדים לתואר שני בפסיכולוגיה קלינית. השאלונים הועברו לכל המועמדים. כולם התקבלו ללימודים. לאחר שנת הלימודים הראשונה נבחן כל מועמד וקיבל ציון על הישגיו בתום שנה זו (Y1). כמו כן בתום השנה העריך המורה את היכולת של כל מועמד להיות פסיכולוג קליני באמצעות ריאיון (Y2). הנח כי כל המדדים היו מהימנים.

שימו לב כי :

X1 – ככל שהציון בשאלון גבוה יותר הרגישות נמוכה יותר.

X2 – ככל שהציון בשאלון גבוה יותר היכולת האמפאטית גבוהה יותר.

Y1 – ככל שהציון גבוה יותר ההישגים גבוהים יותר.

Y2 – ככל שהציון גבוה יותר הערכת המועמד כמתאים להיות פסיכולוג גבוהה יותר.

להלן התוצאות שהתקבלו :

המתאם בין X1 ל-Y1 : 0.4-.

המתאם בין X1 ל-Y2 : 0.3.

המתאם בין X2 ל-Y1 : 0.03.

המתאם בין X2 ל-Y2 : 0.4.

א. לגבי כל מתאם ציין באיזו בדיקת תוקף מדובר.

ב. מה תוכלו לומר על ערכו של שאלון הרגישות לאור נתוני התוקף שלעיל? הסבירו.

ג. מה תוכלו לומר על ערכו של שאלון היכולת האמפאטית לאור נתוני התוקף שלעיל? הסבירו.

תשובות סופיות:

1) ב'.
2) מהימנות ותוקף מבחין טובים, תוקף מתכנס חלש.
3) ב'.
4) ב'.
5) א'.
6) ד'.
7) ב'.
8) ב'.
9) ב'.
10) ראו סרטון.

תרגול מסכם – תוקף ומהימנות:

תרגול:

1) חוקר בנה מבחן לבדיקת כושר מכני (להלן מבחן 1), הכולל 25 פריטים, לכל פריט שתי תשובות אפשריות, אחת נכונה והשנייה לא נכונה.

א. על מנת לבדוק את מהימנות המבחן בנה החוקר נוסח נוסף (להלן מבחן 2) והעביר אותו לאותם אנשים באותו הזמן. החוקר בחן את הקשר בין הציונים שהתקבלו במבחן 1 לבין הציונים שהתקבלו במבחן 2. איזו בדיקה ביצע החוקר ומה מטרתה? הסבר בקצרה.

ב. עוזר המבחן הציע להעביר את מבחן 1 (בנוסף להעברתו הראשונה) לקבוצה אחרת של נבדקים ובמועד אחר.
   i. ציינו איזו בדיקת מהימנות התכוון עוזר המחקר לבדוק?
   ii. האם אמנם ניתן לבדוק מהימנות זו בדרך שהוצעה? הסבירו בקצרה.

ג. החוקר העביר את מבחן 1 לשני סוגי נבדקים: מדגם של מכונאי רכב ומדגם של פקידים.
   i. איזה בדיקה ערך החוקר? בדיקת תוקף או מהימנות?
   ii. מאיזה סוג? הסבר באמצעות מונחי שיטות ותכונות.
   iii. איזה תוצאה תהיה רצויה לו ואיזה לא? הסבר בקצרה.
   iv. מהו סוג המשתנה הנבדק? הסבר בקצרה.

ד. 30 נבדקים שענו על מבחן 1, התקבלו לעבודה כמכונאים ללא התייחסות לציונם במבחן. החוקר בדק האם יש קשר בין הצלחתם במבחן להערכת עבודתם כמכונאים על ידי הממונה הישיר עליהם. להלן התוצאות (בתאים מספר האנשים):

הערכת הממונה	הישגים במבחן		סה"כ
	הצלחה	כישלון	
גרוע	5	5	10
טוב	10	10	20

האם למבחן יש תוקף ניבוי? (על מנת לענות על נשאלה אין צורך בחישוב מעבר לחישוב אחוזים).

2) כחלק ממיון מועמדים להוראה בסמינר למורים, הועבר ל-80 מועמדים מבחן יכולת עמידה מול כיתה. המועמדים נתבקשו לתת שיעור לכיתה למשך חצי שעה, כשבאותה עת צפו בהם שני מורים ותיקים (להלן צופה א' וצופה ב'). כל מורה ותיק נתן באופן בלתי תלוי הערכתו לגבי רמת המועמד על פני ממדים שונים, כמו גם ציון סופי של מידת ההתאמה להוראה על סולם שבין 1 (כלל לא מתאים להוראה) לבין 5 (מאוד מתאים להוראה).

א. החוקר בדק את המתאם בין שני המורים הוותיקים. עבור 85% מהנבדקים התקבלה הסכמה בין שני הבודקים.
   i. האם זו בדיקת תוקף או מהימנות?
   ii. מאיזה סוג? הסבירו בקצרה.
   iii. האם ניתן לומר ש-85% מהשונות האמיתית משוחררת מבעיות הקשורות בזמן? הסבירו.

ב. הנבדקים התבקשו לאחר כחודשיים לתת שיעור נוסף ששוב צפו בהם שני המורים הוותיקים (צופה א' וצופה ב') ונתנו את הערכתם לגבי מידת ההתאמה להוראה. עבור צופה א' נתקבל מתאם של 0.8 בין הערכתו את השיעור הראשון לבין הערכתו את השיעור השני.
   i. האם זוהי בדיקת תוקף או מהימנות?
   ii. מאיזה סוג? הסבירו במונחי תכונות שיטות.

ג. כל המועמדים התקבלו לסמינר בלא התחשבות בהערכת שני הצופים. החוקר בדק את הקשר בין הציון שהתקבל על ידי צופה ב' לבין ממוצע הציונים של המועמד שהתקבל בסיום הסמינר. התקבל מתאם של 0.34.
   i. איזו בדיקה ערך החוקר? בדיקת תוקף או מהימנות?
   ii. מאיזה סוג? הסבירו באמצעות שיטות ותכונות.

ד. החוקר המשיך לעקוב אחר תפקודם של המועמדים בעבודה כמורים. כל מועמד הוערך על תפקודו כמורה על ידי מנהל בית ספרו (בסולם שבין 1-גרוע מאוד ל-7-מצויין). נתקבל מתאם של 0.5 בין ממוצע הציונים של המועמד בסמינר לבין הערכת המנהל (הנח כי כל 80 המועמדים התקבלו לעבודה כמורים).
   i. האם זוהי בדיקת תוקף או מהימנות?
   ii. מאיזה סוג? הסבירו באמצעות שיטות תכונות.
   iii. באיזה סולם מדידה נתונה מידת ההערכה להתאמה להוראה?
   iv. האם ניתן לערוך טרנספורמציה על ציוני סולם זה של הוספת קבוע חיובי? הסבירו בקצרה.

3) חוקר בנה שאלון חדש שמטרתו הייתה לבדוק את אהבת קריאת הספרים של האדם. השאלון כלל 25 פריטים, כשסולם התשובות נע בין 1 ל-5 לכל פריט (1 - חוסר עניין מוחלט בקריאת ספרים ו-5 - אהבה רבה של קריאת ספרים). השאלון הועבר למדגם של 100 נבדקים.

א. בבדיקה שערך החוקר התקבלה α של קרונבך 0.8.
   i. איזה בדיקה ערך החוקר?
   ii. על סמך הבדיקה: מה אחוז השונות הנובע מטעויות?
   iii. מאיזה מקור לטעויות משוחררת בדיקה זו?

ב. מהו המקור האפשרי לשונות אמיתית בציוני הנבדקים בשאלון אהבת הקריאה?

ג. החוקר העביר את השאלון למדגם של נשים ולמדגם של גברים. לא נמצא הבדל בציוני שתי הקבוצות.
   i. איזו בדיקה ערך החוקר?
   ii. מה מעידה התוצאה על טיבו של השאלון?

ד. החוקר העביר את השאלון לאותם נבדקים לאחר חמישה חודשים. נמצא מתאם של 0.9 בין הציונים של ההעברה הראשונה לציונים של ההעברה השנייה.
   i. איזו בדיקה ערך החוקר? תוקף או מהימנות?
   ii. מאיזה סוג? הסבר במונחי תכונות שיטות.
   iii. האם ניתן לומר על סמך התוצאות ש-90% מהשונות היא שונות אמיתית המשוחררת בתלות בתוכן?

ה. בבדיקה נוספת נבדק המתאם בין ציון בשאלון הבודק את מס׳ הספרים שקרא הנבדק בשנה האחרונה לבין ציונו בשאלון אהבת הקריאה.
   i. איזו בדיקה ערך החוקר-תוקף או מהימנות?
   ii. מאיזה סוג. הסבר במונחי תכונות שיטות.

ו. בבדיקה נוספת הועבר השאלון אהבת הקריאה לקבוצת נבדקים הפוקדים מדי שנה את אירועי שבוע הספר לבין קבוצת נבדקים שמעולם לא ביקרה בשבוע הספר. לא נמצא הבדל מובהק בין שתי הקבוצות בציון בשאלון.
   i. איזו בדיקה בוצעה-תוקף או מהימנות?
   ii. מאיזה סוג? הסבר במונחי תכונות שיטות.
   iii. מה מעידה תוצאת הבדיקה על טיבו של השאלון?

ז. נמצא מתאם של 0.5 בין ציון במבחן הבוחן את הרמה המילולית של הנבדק לבין ציון בשאלון אהבת הקריאה.
   i. איזה בדיקה נערכה-תוקף או מהימנות?
   ii. מאיזה סוג? הסבר במונחי תכונות שיטות.
   iii. מה מעידה התוצאה על טיבו של השאלון?

תשובות סופיות:

1) א. אקוויולנטיות, מטרה: זיהוי רגישות הכלי לייצוגיות הדגימה של פרטי התוכן.
ב.i. מבחן חוזר. ii. לא. ג.i. תוקף. ii. מבנה.
iii. ראו סרטון. iv. מנה. ד. לא.

2) א.i. מהימנות. ii. בין שופטים. iii. לא. ב.i. מהימנות.
ii. מבחן חוזר. ג.i. תוקף. ii. תלוי קריטריון – ניבוי.
ד.i. תוקף. ii. תלוי קריטריון – ניבוי. iii. סדר. iv. כן.

3) א.i. עקיבות פנימית. ii. 20%. iii. ייצוגיות הדגימה של פרטי התוכן.
ב. ההבדל בין האנשים בתכונה הנמדדת. ג.i. תוקף מבנה. ii. תמיכה.
ד.i. מהימנות. ii. מבחן חוזר. iii. לא. ה.i. תוקף.
ii. מתכנס. ו.i תוקף. ii. מבנה. iii. לא תומכת.
ז.i. תוקף. ii. מבחין. iii. נמוך.