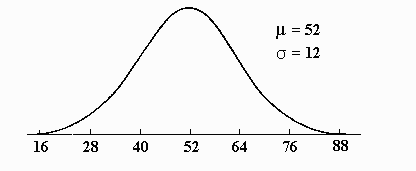
המרכז האקדמי לב

**סטטיסטיקה למהנדסים**



# מאת

ד"ר אלכסנדר קליין

מבוא: סולמות מדידה

1. משתנה שמי: הערכים השונים הם מזהים בלבד, ויתכן שלא יהיו מספרים אלא מלים בלבד. דוגמה: שמות של אנשים, מספרי טלפון.
2. משתנה סידורי: הערכים השונים הם מספריים, והם מצביעים אך ורק על סדר מסויים. דוגמה: תוארים אקדמאיים (ראשון, שני, שלישי).
3. משתנה רווחי: יש חשיבות לרווחים שבין שני ערכים שונים, אבל לא לערכים עצמם. דוגמה: טמפרטורה.
4. משתנה מנה: לערכים המספרים יש משמעות. דוגמה: גובה המשכורת.

**פרק ראשון: הצגת נתונים בטבלאות ועקומות**

1. דוגמה מספרית: מדגם של 31 יום, הטמפרטורה בחודש אוגוסט בשעה 12.00:

30,29,28,27,29,30,30,27,26,26,25,26,30,32,30,30,30,31,34,33,27,28,29,30,30,29,30,31,27,29,29

השאלה: איך לארגן ולהציג את הנתונים, כך שנוציא מהם את המידע המרבי?

#### הגדרות

* + *משתנה*: גודל העשוי לקבל ערכים שונים. דוגמה: טמפרטורה בשעה 12.00 באשקלון.
  + *שכיחות*: כמות ערכים של משתנה הנמצאים בטווח נתון. דוגמה: השכיחות של הטמפרטורה 30 היא 10. מסמנים אותה על ידי f(x) (frequency).

ניתן להציג את התפלגות השכיחויות, כדלהלן:

|  |  |
| --- | --- |
| f(x) | x |
| 1 | 25 |
| 3 | 26 |
| 4 | 27 |
| 2 | 28 |
| 6 | 29 |
| 10 | 30 |
| 2 | 31 |
| 1 | 32 |
| 1 | 33 |
| 1 | 34 |

## או תוך שימוש בנתונים מקובצים:

|  |  |
| --- | --- |
| f(x) | x |
| 4 | 25-26 |
| 6 | 27-28 |
| 16 | 29-30 |
| 3 | 31-32 |
| 2 | 33-34 |

לשם כך, יש להקפיד על שני תנאים הכרחיים:

* + המחלקות חייבות להיות זרות
  + החלוקה חייבת להיות ממצה

בבניית טבלת שכיחויות עלינו למצוא את שביל הזהב בין הקיבוץ המזערי, בו ריבוי הפרטים מקשה עדיין על ההתרשמות, לבין הקיבוץ המרבי, בו ניטשטשו רוב הפרטים החשובים.

* + לכל מחלקה יש *גבול עליון* *וגבול תחתון*
  + *גבולות מדומים*: מדובר בגבולות מדומים כאשר הגבול העליון של מחלקה איננו מתלכד עם הגבול התחתון של המחלקה שמתחתיה. אם זה כן מתלכד, מדובר בגבולות *אמיתיים*.

מטבלה בעלת גבולות מדומים ניתן לקבל טבלה בעלת גבולות אמיתיים, על ידי חלוקת הרווח בין המחלקות לשניים.

דוגמה:

|  |  |
| --- | --- |
| f(x) | x |
| 4 | 24.5-26.5 |
| 6 | 26.5-28.5 |
| 16 | 28.5-30.5 |
| 3 | 30.5-32.5 |
| 2 | 32.5-34.5 |

אין סתירה לכלל הראשון, כאשר הגבול המשותף איננו מופיע בנתונים.

* + *נקודת אמצע* של מחלקה: ערך הממוצע בין הגבול העליון לבין הגבול התחתון שלה. דוגמה: 33.5 עבור המחלקה העליונה.
  + *רוחב המחלקה*: ההפרש בין גבול עליון אמיתי לבין גבול תחתון אמיתי. דוגמה: 2.

בדרך כלל, נהוג לקבוע רוחב אחיד לכל המחלקות, אך אין זה הכרחי. דוגמה: מספר שעות שהסטודנטים מקדישים ללימודים כל שבוע:

|  |  |
| --- | --- |
| f(x) | x |
| 1 | 9> |
| 3 | 9-12 |
| 11 | 13-16 |
| 29 | 17-20 |
| 2 | 20< |

#### שכיחות יחסית ומצטברת

שכיחות יחסית: , שכיחות מצטברת: . שכיחות יחסית מצטברת: 

דוגמה:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| F(x)/n | F(x) | f(x)/n | f(x) | x |
| 4/31 | 4 | 4/31 | 4 | 24.5-26.5 |
| 10/31 | 10 | 6/31 | 6 | 26.5-28.5 |
| 26/31 | 26 | 16/31 | 16 | 28.5-30.5 |
| 29/31 | 29 | 3/31 | 3 | 30.5-32.5 |
| 1 | 31 | 2/31 | 2 | 32.5-34.5 |

פלט SPSS:

****

#### דיאגרמת מקלות ומצולע

* + דיאגרמת מקלות הוא תרשים במערכת צירים, בו הציר האופקי מייצג את ציר המשתנה x, (לוקחים את אמצעי המחלקות), ואילו הציר האנכי מייצג את ציר השכיחויות f(x) (או שכיחות יחסית, וכו').
  + ****
  + מצולע מתקבל מדיאגרמת מקלות על ידי מתיחת קווים בין ראשי המקלות השונים.

דוגמה:



1. הסטוגרמה

יש לעבור לגבולות אמיתיים כדי לבנות הסטוגרמה, כדלהלן:



אם רוחב המחלקה איננו אחיד, גובה כל מלבן נקבע על ידי חילוק השכיחות ברוחב המחלקה. רוחב המלבן הוא בכל מקרה רוחב המחלקה.

דוגמה:

|  |  |  |  |
| --- | --- | --- | --- |
| גובה | רוחב | f(x) | x |
| 2 | 1 | 2 | 3-4 |
| 4 | 3 | 12 | 4-7 |
| 6 | 2 | 12 | 7-9 |
| 4 | 1 | 4 | 9-10 |
| 2 | 3 | 6 | 10-13 |

**פרק שני: מדדים למיקום מרכזי**

#### ערכים בודדים

*שכיח (Mode)*: אותו ערך המופיע בשכיחות הגבוהה ביותר. דוגמה: 30.

*אמצע הטווח (Midrange)*: הממוצע בין התצפית הגדולה ביותר והקטנה ביותר. דוגמה: 

*חציון (Median)*: כאשר נתונה רשימת תצפיות שסודרו בסדר עולה, זה ערך התצפית המרכזית אם מספר התצפיות אי-זוגי. אם מספרם זוגי, הוא ממוצע ערכיהן של שתי התצפיות המרכזיות. דוגמה: 29.

*ממוצע (Mean)*: . דוגמה: 29.09

#### נתונים מקובצים

*שכיח*: בתנאי שרוחב המחלקה אחיד, הוא נקודת האמצע של המחלקה השכיחה ביותר. יתכן מספר שכיחים. דוגמה: 29.5.

*אמצע הטווח*: הוא הממוצע בין הגבול העליון של המחלקה העליונה לבין הגבול התחתון של המחלקה התחתונה.

דוגמה: 

*חציון*:



כאשר:

n: מספרם הכולל של התצפיות.

 : המחלקה בה נמצא החציון.

: המחלקה הקודמת ל-

: השכיחות המצטברת עד למחלקה 

: שכיחות המחלקה 

: גבול תחתון אמיתי של המחלקה שבה נמצא החציון

: גבול עליון אמיתי של המחלקה שבה נמצא החציון

דוגמה:



הערה: אם מספר התצפיות הינו זוגי ושתי תצפיות המרכזיות נמצאות בשתי מחלקות נפרדות, אזי החציון הוא הגבול המשותף לשתי המחלקות.

ובאופן יותר כללי, להלן ההגדרה של המאון p (או אחוזון p, percentile), כאשר הוא המחלקה שבה נמצא המאון שמדובר עליו:



*ממוצע*:



כאשר x הוא האמצע של כל מחלקה. דוגמה: 29.048

****

הערה: החציון של הנתונים המקובצים איננו מדויק.

#### השוואה בין המדדים השונים

אמצע הטווח (במידה גדולה) והממוצע (במידה פחותה) מושפעים מערכים קיצוניים, ואילו השכיח והחציון אינם מושפעים מהם.

השכיח והחציון נשענים יותר על שכיחות התצפיות השונות, ואילו הממוצע ואמצע הטווח יותר על ערכיהן.

הממוצע הוא בדרך כלל ומן הסתם המדד המשקף ביותר ולכן הוא המדד השימושי ביותר.

**פרק שלישי: מדדים לפיזור**

אין כאן הבחנה בין ערכים בודדים לנתונים מקובצים.

*טווח (Range)*: ההפרש בין הערך הגבוה ביותר והערך הנמוך ביותר. דוגמה: בערכים הבודדים: 34-25=9

*טווח בין-רבעוני*: , במטרה לנטרל את השפעתם של הנתונים החריגים.

*ממוצע הסטיות המוחלטות מהממוצע*:

הגיוני, אבל לא נוח, ולכן לא שימושי.

*שונות (Variance)*:

1. ערכים בודדים:



דוגמה: 4.15.

1. נתונים מקובצים:



דוגמה: 4.05.

*סטיית תקן(Standard Deviation)*: 

משפט: בנתונים בודדים: 

ציון תקן (לשם השוואה בין שתי אוכלוסיות שונות):

**פרק רביעי: אמידה נקודתית**

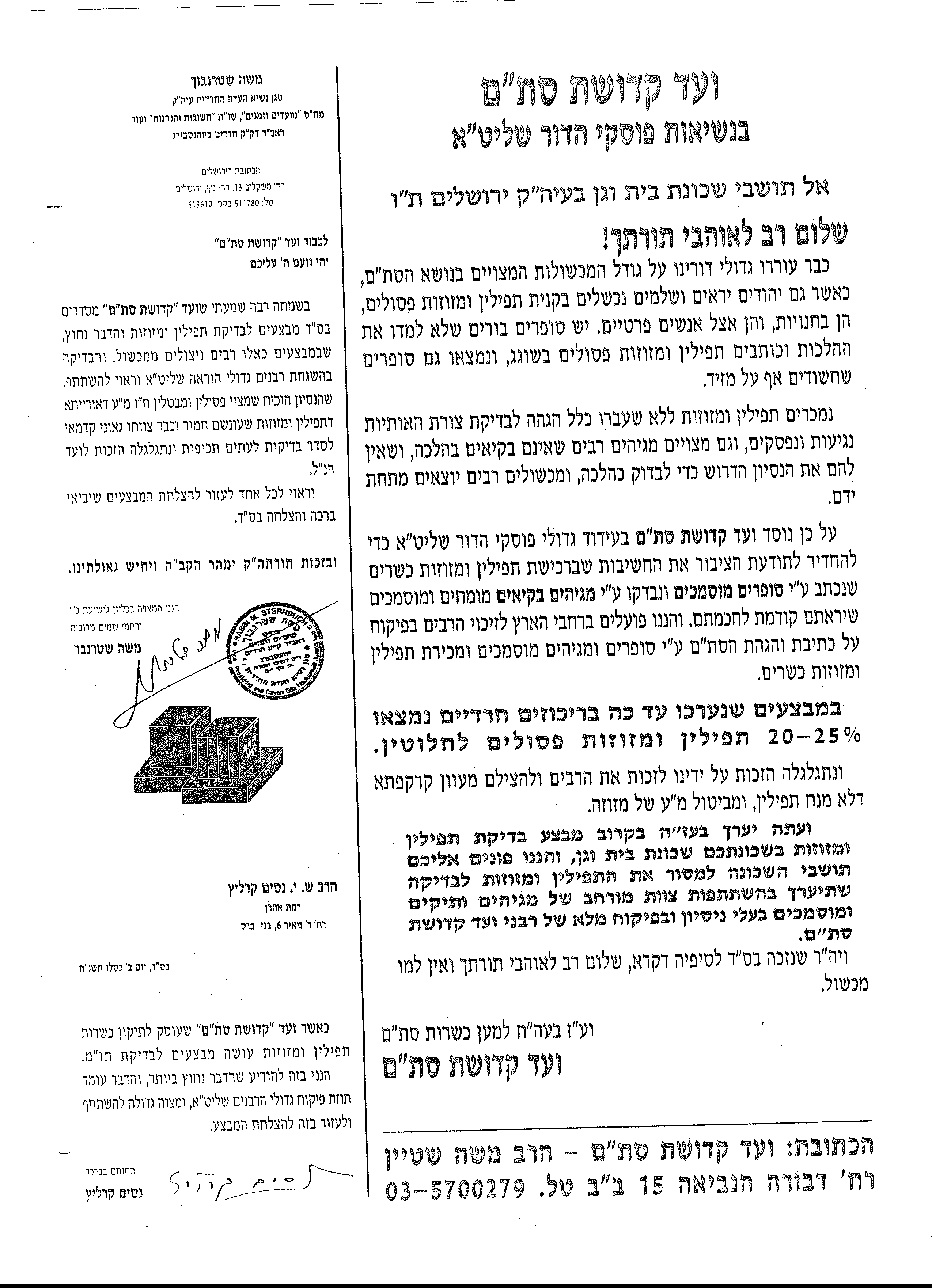
1. **דגימה מקרית:**
   1. מושגים בסיסיים
      1. אוכלוסיה: כלל הפריטים לגביהם אנו שואלים שאלה מסוימת
      2. פרמטר: גודל קבוע המאפיין את האוכלוסייה (למשל, p, σ, μ)
      3. מדגם מקרי: תת-אוכלוסיה שעליה בודקים את נושא השאלה, כאשר יש הסתברות שווה לכל איבר באוכלוסיה להופיע במדגם
      4. אמידה: על סמך תוצאות המדגם, חישוב ערך משוער לפרמטר המבוקש.
   2. ייצוג מתמטי של המדגם (X1,...,Xn)

* לכל Xi אותה פונקצית הסתברות
* כל ה-Xi הם בלתי תלויים

1. **אמידה:**
   1. סטטיסטי: 

דוגמה: 

* 1. אומד (Estimator):  בתור משתנה מקרי: 
  2. אומדן (Estimation): תצפית של :  על סמך מדגם מסוים

****

1. **הטייה:**
   1. הגדרה: 

האומד  הינו חסר הטייה אם 

* 1. אומד חסר הטייה עבור התוחלת:

משפט:  הוא אומד חסר הטייה עבור μ.

הוכחה: 

מ.ש.ל.

* 1. אומד חסר הטייה עבור השונות:

משפט: אם μ ידוע, אזי  הוא אומד חסר הטייה עבור σ2.

הוכחה: 

מ.ש.ל.

משפט: אם μ אינו ידוע,אזי  הוא אומד חסר הטייה עבור σ2.

הוכחה:







מ.ש.ל.

הערה: 

* 1. אומד חסר הטייה עבור הפרמטר p של התפלגות בינומית:

נתון: k – מספר הצלחות

n – מספר ניסויים

אם כן, האומד המתקבל על הדעת הוא:  , כאשר  .



לכן  הוא אומד חסר הטייה.

* 1. אומד חסר הטייה עבור הפרמטר θ של התפלגות אחידה:

נניח 

ניקח עכשיו: .



 הוא אומד חסר הטייה.

* 1. אומד חסר הטייה בצורה אסימפטוטית:

א) הגדרה: 

ב) דוגמה:

1. **שיטת הנראות המרבית:**
   1. נראות של המדגם (X1,...,Xn) 

(במקרה בדיד, לוקחים עבור : )

* 1. שיטת הנראות המרבית

הגדרה: אומדן בעל נראות מרבית  מקיים את התנאי הבא:

דרך החישוב:





* בעקרון, צריכים לבדוק שהנגזרת השנייה שלילית.

דוגמה:  



1. **אומד מתכנס:**

הגדרה: אומד נקרא מתכנס אם השונות שלו שואפת לאפס 

דוגמה:  



1. **השוואה בין אומדים שונים:**

יהיו נתונים שני אומדים חסרי הטייה. נעדיף את האומד בעל השונות הנמוכה מביניהם.

דוגמה:



יוצא כי  עדיף על פני  אם מתקיים 

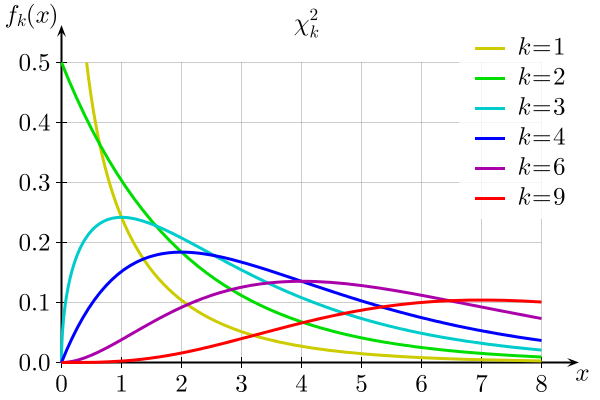
פרק חמישי: התפלגויות של סטטיסטיים

1. **התפלגות חי-בריבוע**

משפט: יהיו נתונים k משתנים נורמליים תקניים בלתי-תלויים , אזי, המשתנה המקרי  מתפלג על פי פונקצית הצפיפות הבאה:

.

נאמר כי משתנה זה מתפלג על פי התפלגות חי-בריבוע, עם k דרגות חופש:.



תכונות:

א) 

הוכחה:



מ.ש.ל.

ב) 

משפט: יהיו נתונים p משתנים בלתי-תלויים המתפלגים על פי התפלגות חי-בריבוע , אזי .

משפט: יהי נתון מדגם אקראי , כאשר , אזי מתקבל .

מכאן שאומד השונות הוא מתכנס.

הוכחה:



שימוש בטבלאות: ברצוננו למצוא את a כך ש- . התשובה היא 21.026.

קירוב נורמלי: אם k>30, אזי הכמות  מתפלג על פי התפלגות נורמלית תקנית. מכאן נובע:





אם כן, למשל: ברצוננו למצוא את a כך ש-. מכאן נובע



1. **התפלגות t**

משפט: יהי נתון Z~N(0,1), ו-, Z ו-V בלתי תלויים. אזי המשתנה המקרי  מתפלג על פי פונקצית הצפיפות הבאה:



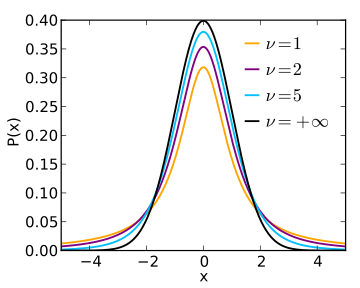
נאמר כי .

משפט: יהי נתון מדגם אקראי , כאשר , , אזי .

הוכחה:

**

מ.ש.ל.



שימוש בטבלאות: ברצוננו למצוא את a כך ש-. התשובה היא 2.08596.

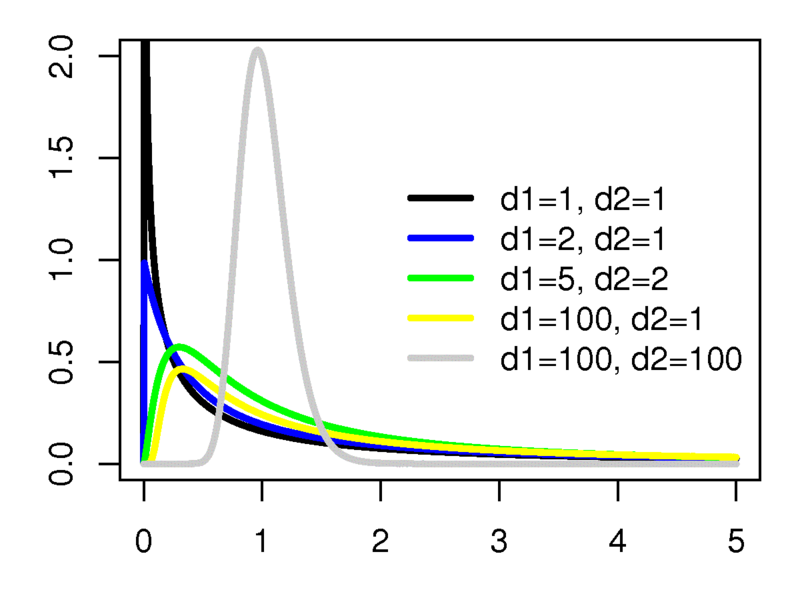
1. **התפלגות F**

משפט: יהיו נתונים שני משתנים מקריים בלתי-תלויים , , אזי המשתנה המקרי  מתפלג על פי פונקצית הצפיפות הבאה:



נאמר כי .

תכונה חשובה: .



משפט: יהיו נתונים שני מדגמים נורמליים בלתי תלויים . אזי, מתקבל:



הוכחה:

, 

והם בלתי תלויים.

אם כן:



מ.ש.ל.

שימוש בטבלאות: ברצוננו למצוא את a כך ש- . התשובה היא 5.39.

**פרק שישי: רווחי סמך**

1. **רווח-סמך עבור μ כאשר σ ידוע:**

משפט: אם  או , אזי :



הוכחה: אם X~N או , אזי .



0

מ.ש.ל.

ולמעשה:



דוגמה: . בנה רווח סמך ברמת בטחון 0.95 עבור μ, כאשר התקבל , n=10.



1. **רווח-סמך עבור μ כאשר σ אינו ידוע:**
   1. n גדול ():

אז אפשר להחליף את σ ב-s, כאשר, ולהשתמש בנוסחה הנ"ל, כאשר:





* 1. n קטן ():
     1. X אינו מתפלג נורמלית: אין פתרון.
     2. X מתפלג נורמלית. אזי:



ולמעשה:



דוגמה: :



1. **רווח-סמך עבור פרופורציה:**
   1. אומד עבור p



ובכן  הוא חסר הטייה



ובכן, אם n גדול, לפי משפט הגבול המרכזי:



* 1. רווח סמך עבור p (אם n גדול, למעשה )



ולמעשה:



דוגמה: מתוך 100 לידות, נולדו 48 בנים ו-52 בנות. אם כן, הרווח-סמך עבור ההסתברות ללידת בן ברמת בטחון 0.95 יהיה:

 היינו  היינו 

**פרק שביעי: בדיקת השערות**

1. **מבוא:**

דוגמה: מוכר נורות טוען שבממוצע, משך חיים של נורותיו הינו 200 שעות. אגודת הצרכנים רוצה לבדוק אם טענתו נכונה: היא מוציאה מדגם של 50 נורות ומחשבת את הממוצע של המדגם. היא תחליט שלא לקבל טענתו של המוכר אם יש מרחק רב בין הממוצע שהתקבל לבין 200. אבל מה הגבול?

1. **הגדרות יסודיות:**

* השערת אפס *H0*: הטענה הפשוטה 
* השערה אלטרנטיבית *H1*: 

ההשערה האלטרנטיבית יכולה להיות או חד-צדדית, או דו-צדדית.

* סוגי טעויות:

**ההחלטה**

|  |  |  |
| --- | --- | --- |
| **המצב**  **האמיתי** | *H0* | *H1* |
| *H0* | אין  טעות | I |
| *H1* | II | אין  טעות |

α: הסתברות לטעות מסוג I

β: הסתברות לטעות מסוג II

1-β: עוצמת המבחן

* אזור קריטי (או אזור הדחייה): האזור שאם בו נופל הערך שהתקבל ע"י המדגם, מחליטים לדחות את *H0*. ובכן, α הוא ההסתברות של האזור הקריטי, בתנאי ש- *H0*היא נכונה.
* מובהקות (P-Value): ההסתברות לקבל ערך "קיצוני" יותר מהסטטיסטי שהתקבל בפועל, בהנחה שהשערת האפס היא נכונה. אם המובהקות קטנה מרמת המובהקות שהחליטו עליה, אזי דוחים את השערת האפס.



1. **הפעלת הבדיקה:**

* קביעת ההשערות *H0* ו- *H1*
* קביעת המקרה שבו נמצאים
* קביעת האזור הקריטי
* חישוב הסטטיסטי
* בדיקת מיקום הסטטיסטי ומסקנה

1. **סוגים שונים של מקרים:**

**4.1: השערה פשוטה מול השערה פשוטה**

במקרה הזה, ניתן לחשב את וגם את , בהתאם למנגנון ההחלטה.

דוגמה:

נתון כי זמן ההמתנה לרכבת הקלה מתפלג התפלגות אחידה רציפה, היינו . יש התלבטות אם הרכבת מגיעה כל 10 דקות () או כל חצי-שעה (). מנגנון ההחלטה הוא כדלהלן: אתה מגיע פעם אחת לתחנה. אם חיכית 10 דקות או פחות עד הגעת הרכבת, תסיק כי , ואילו אם חיכית יותר מ-10 דקות, תסיק כי . חשב את ו-.

פתרון:

**4.2: השערה פשוטה מול השערה מורכבת**

➀

* 1. מבחן על μ , X מתפלג נורמלית או n>30 , σ ידוע

   ➀

אזור קריטי:   

הסבר: על פי השערת האפס, .

דוגמה: דיקן האוניברסיטה טוען שרמת הסטודנטים בלשון העברית היא 67.5 (תוחלת). אגודת הסטודנטים טוענת שאין פרט זה נכון. לשם בדיקת טענתו, נלקח מדגם של 100 סטודנטים, והתקבל ממוצע של 68.5. ידועה סטיית התקן, על-פי ניסיון העבר, והיא שווה 5. מה ניתן להסיק ברמת מובהקות של 0.05?

 נכון! מסקנה: דוחים את *H0*, הדיקן לא דייק.

חישוב המובהקות: .

➁

* 1. מבחן על μ , X מתפלג נורמלית, σ אינו ידוע
     1. אם , כמו ב-4א', רק מחליפים σ ב-*s*.
     2. אם , פועלים כדלהלן:

   ➁

אזור קריטי:   

הסבר: על פי השערת האפס, .

דוגמה: יצרן תמרוקים רוצה לבדוק אם תוחלת משקלן של השפופרות הממולאות במפעלו היא אמנם 103 גרם. התקבלו הנתונים הבאים (ניתן להניח התפלגות נורמאלית): 96, 102, 98, 102, 104, 98, 100. מה ניתן להסיק, ברמת מובהקות 0.02?

 לא נכון! מסקנה: לא דוחים את *H0*, התוחלת איננה שונה מ-103 בצורה מובהקת.



* 1. מבחן על , X1 ו-X2 מתפלגים נורמלית או ,  ידועים (מדגמים בלתי תלויים)

   ➂

אזור קריטי:   

הסבר: על פי השערת האפס,.

דוגמה: חוקר מעוניין לבדוק אם יש הבדל מבחינת התוחלת בין יכולתם של גברים ליכולתן של נשים במבחן "יצירת מילים". הוא נתן שש אותיות שונות לקבוצה המונה 114 גברים ו-175 נשים, וכל נבדק נתבקש להרכיב במשך 5 דקות מילים רבות ככל האפשר מאותיות אלו. נניח שמניסיון קודם ידוע שסטיית התקן של היכולת במבחנים כאלה היא 6 אצל הגברים ו-5 אצל הנשים. ממוצע מספר המילים של הנשים היה 21 ושל הגברים 19.7 ().

 לא נכון! מסקנה: לא דוחים את *H0*, אין הבדל מובהק בין שני המגדרים.

חישוב המובהקות: .

* 1. מבחן על , X1 ו-X2 מתפלגים נורמלית,  לא ידועים
     1. אם , כמו ב-4ג', רק מחליפים σ ב-*s*.
     2. אם , פועלים כדלהלן:

   ➃

אזור קריטי:   

הסבר: ראו דף תרגילים 6 שאלה 11.

דוגמה: חוקר רוצה לבדוק אם קיים הבדל מובהק בין הישגים של בנים לבין הישגים של בנות במבחן בסטטיסטיקה. ידוע שציוני המבחן מתפלגים נורמאלית (מניחים שסטיות התקן שוות, ).

 נכון! מסקנה: דוחים את *H0*, קיים הבדל מובהק בין המגדרים.

****

****

* 1. מבחן על p,  וגם 

   ➅

אזור קריטי:   

הסבר: על פי השערת האפס,.

דוגמה: שחקן טוען שהסתברות להוציא "6" על קובייה מסוימת קטנה מ-. התקבל "6" שמונה פעמים ב-100 הטלות. מי צודק, ברמת מובהקות 0.05?

 נכון! מסקנה: דוחים את *H0*, הקובייה איננה הוגנת.

חישוב המובהקות: 

ו. מבחן על הפרש שונויות, X1 ו-X2 מתפלגים נורמלית

  ➈

אזור קריטי:  



הערה: אם המבחן חד-צדדי, לוקחים כ- את המשתנה בעל השונות המרבית לפי .

הסבר: על פי השערת האפס, .

דוגמה: בדוגמה שבסעיף ד', בכוונתנו לבחון אם אמנם ניתן להניח שאין הבדל מובהק בין שתי השונויות, ברמת מובהקות 0.1.

 נכון! מסקנה: לא דוחים את *H0*, אין הבדל מובהק בין שתי השונויות.

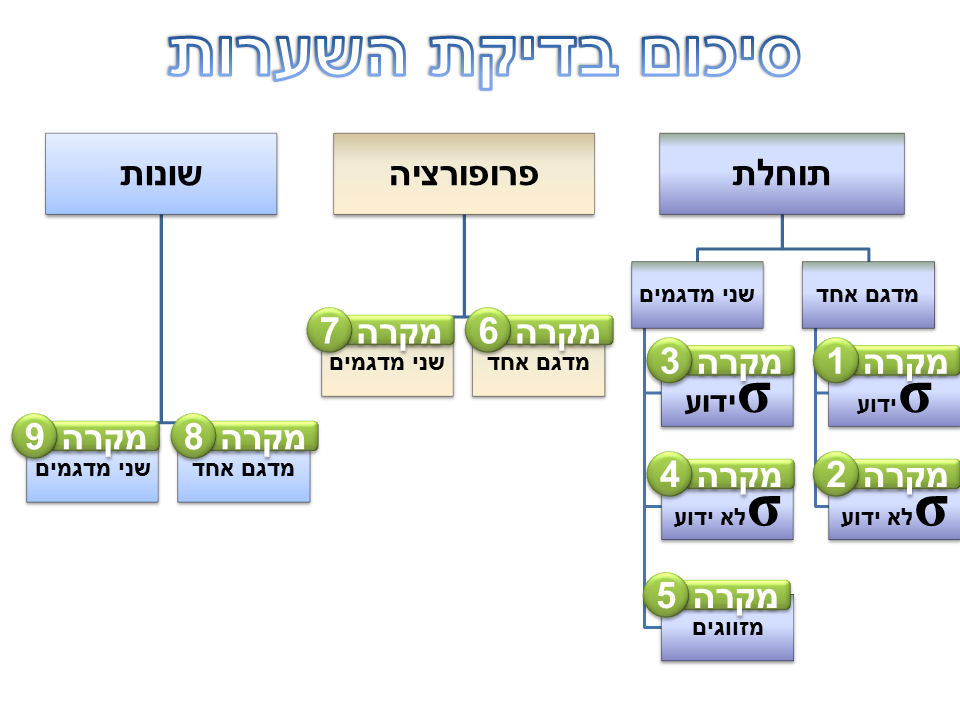
הערה: 

**סיכום בדיקת השערות ורווחי סמך**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | סטטיסטי/מבחן |  | אזור דחייה | רווח סמך |
|  |  | ידוע ו  או |  |  | של : |
|  |  |
|  |  |
|  |  | לא ידוע, ו- |  |  | של : |
|  |  |
|  |  |
|  |  | ידועים  או |  |  | של : |
|  |  |
|  |  |
|  |  | לא ידועים      וכן |  |  | של : |
|  |  |
|  |  |
|  |  | מדגמים מזווגים |  |  | של : |
|  |  |
|  |  |
|  |  |  |  |  | של : |
|  |  |
|  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | סטטיסטי/מבחן |  | אזור דחייה | רווח סמך |
|  |  | , אזי לוקחים כמכנה:  , כאשר |  |  | של : |
|  |  |
|  |  |
|  |  |  |  |  | של :  בין  לבין |
|  |  |
|  | או |
|  |  |  |  |  | של :  בין  לבין |
|  | או |





באדיבות דוד מנסבך**פרק שמיני: מבחן לבדיקת טיב התאמה**

1. **התאמה להתפלגות:**

המטרה: לבדוק אם אוכלוסיה מסוימת מתפלגת לפי התפלגות נתונה, וזו השערת האפס:

* Oi : מספר תצפיות ששויכו לטווח *i*
* Ei : מספר תצפיות אמורות להשתייך לאותו טווח אם ההשערה נכונה
* *k* טווחים, *n* תצפיות בסך הכל.



הערה: Ei חייב להיות גדול או שווה ל-5, אחרת מחברים טווחים.

אם , דוחים את ההשערה.

* *t* : מספר הכמויות שצריכים להסיק מן הנתונים כדי לחשב את ה-Ei

דוגמה: הטלת קובייה 120 פעם. רוצים לבחון אם הקובייה מאוזנת.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 6 | 5 | 4 | 3 | 2 | 1 |  |
| 24 | 19 | 18 | 17 | 22 | 20 |  |
| 20 | 20 | 20 | 20 | 20 | 20 |  |
| 16 | 1 | 4 | 9 | 4 | 0 |  |
|  |  |  |  |  | 0 |  |







1.7>11.07: לא נכון! מסקנה: לא ניתן לדחות את ההשערה כי הקובייה מאוזנת.

****

****

1. **לוחות תלות הדדית:**

המטרה: לבדוק אם קיימת תלות בין שני משתנים מקריים איכותיים. השערת האפס היא שאין תלות בין המשתנים.



אם , מסיקים שיש תלות

הסבר: ,



דוגמה: האם קיימת תלות מובהקת, ברמת מובהקות 0.05, בין דת ואזור מגורים בארצות הברית?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **פרוטסטנטים** | **יהודים** | **קתוליים** | **סה"כ** |
| **חוף מזרחי** | (201.6)  182 | (187.8)  203 | (210.6)  215 | 600 |
| **חוף מערבי** | (134.4)  154 | (125.2)  110 | (140.4)  136 | 400 |
| **סה"כ** | 336 | 313 | 351 | 1000 |







 מסיקים שיש תלות מובהקת.

****

****

**פרק תשיעי: ניתוח שונות**

1. **מבוא:**

נתונות *k* אוכלוסיות. רוצים לבחון אם הן זהות מבחינה מסוימת. זאת אומרת אם למשתנה מסוים יש אותה תוחלת בכל *k* האוכלוסיות:



 כך ש-

דוגמאות:

* השוואת יבול של חיטה מסוגים שונים
* השוואת שיטות הוראה שונות
* השוואה בין טיפולים שונים

הנתונים:

 מספר אחיד (n) של תצפיות בכל אוכלוסיה.

1. **המודל:**



זאת אומרת שדורשים כי:

* Y יתפלג נורמאלית
* השונות של Y צריכה להיות שווה בכל האוכלוסיות (הנחת ההומוסקדסטיסיות)

ולכן אפשר לכתוב:



נסמן:  





SST = SSW + SSB

(Total) (Within) (Between)

אם H0 נכונה, אזי:  מתפלג לפי .

ודוחים את H0 אם: 

1. **נוסחאות לחישוב:**







1. **דוגמה:**

כדי לבחון את התועלת של 4 דשנים שונים, השתמשו בכל אחד ב-5 חלקות, והתקבלו היבולים הבאים (בק"ג). האם יש הבדל מובהק בין הדשנים השונים, ברמת מובהקות 0.05?

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **סה"כ** | **5** | **4** | **3** | **2** | **1** | **חלק**  **דשן** |
| 374 | 80 | 75 | 64 | 73 | 82 | **A** |
| 340 | 60 | 76 | 63 | 84 | 57 | **B** |
| 338 | 75 | 70 | 65 | 45 | 83 | **C** |
| 371 | 80 | 74 | 75 | 74 | 68 | **D** |
| 1423 |  | | | | | |









 אין הבדל מובהק בין הדשנים השונים.

****

|  |  |  |  |
| --- | --- | --- | --- |
| **Test of Homogeneity of Variances** | | | |
| yevul | | | |
| Levene Statistic | df1 | df2 | Sig. |
| 2.139 | 3 | 16 | .135 |

****

1. **הכללה:**

אם אין מספר אחיד של תצפיות בכל אוכלוסיה, אלא 







אם , אזי דוחים את השערת האפס, היינו מסיקים שקיים הבדל מובהק בין תוחלות הקבוצות השונות.

**פרק עשירי: רגרסיה**

1. **מבוא:**

משמעותו של קשר סטטיסטי

מחפשים דרך לחשב את העוצמה ואת הכוון של הקשר בין שני משתנים כמותיים.

דוגמאות:

* טמפרטורה ממוצעת – תנובת עצי פרי.
* מנת משכל של האב – מנת משכל של הבן
* מספר רופאים – מספר ימי מחלה

אם מתגלה קשר בין שני המשתנים X ו-Y, אזי קיימות שלוש אפשרויות:

* X משפיע על Y
* Y משפיע על X
* משתנה שלישי משפיע על X ו-Y.

מקדם המתאם של פירסון (Pearson Coefficient of Correlation)



תכונות:

1. אם |r| קרוב ל-1, זה סימן כי עוצמת הקשר גבוהה. בכל מקרה, מתקיים .
2. אם r חיובי, הקשר ישר, ולא, הקשר הפוך.
3. אם מפעילים על שני המשתנים X ו-Y טרנספורמציה לינארית, מקדם המתאם לא ישתנה, אלא אם כן המקדמים של X ו-Y הם בעלי סימנים הפוכים, שאז המקדם ישנה סימן. זאת אומרת: r(aX+b,cY+d)=r(x,y) בתנאי ש-ac>0, ו- r(aX+b,cY+d)=-r(x,y)אם ac<0.

דוגמה:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| גובה: X | 160 | 160 | 120 | 200 |
| משקל: Y | 75 | 90 | 60 | 75 |



ניתן לבחון את ההשערה לעומת ההשערה בעזרת הטבלה המיועדת לכך (הטבלה האחרונה בחוברת זו), על פי הכלל הבא: אם מקדם המתאם המחושב עולה על הערך המופיע בטבלה, הוא מובהק. במקרה שלנו, המקדם אינו מובהק (אפילו ברמת מובהקות 0.1 = ).

****

מטרות הרגרסיה:

* קשר סטטיסטי בין שני משתנים (באיזו מידה X "מסביר" את Y).
* חיזוי משתנה אחד על סמך משתנה אחר.

דוגמה: משקל וגובה.

1. **רגרסיה ליניארית פשוטה:**

המודל: 



כדי לאמוד את a ואת b, מגדירים: 

ומחפשים a ו-b כך ש-S יהיה מזערי.

מתקבלות הנוסחאות הבאות:  

דוגמה: ציון במתמטיקה ובסטטיסטיקה

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y | 80 | 80 | 60 | 60 |
| X | 60 | 75 | 50 | 40 |





1. **בדיקת מובהקות הרגרסיה:**





SST = SSE + SSR

SST (Total): סה"כ הפיזור.

SSE (Error): הפיזור הבלתי מוסבר ע"י הרגרסיה.

SSR (Regression): הפיזור המוסבר ע"י הרגרסיה.







אם  דוחים את 

נוסחאות לחישוב:







דוגמה (הנ"ל):



שימוש ב-



דוגמה: 

דרך נוספת לחשב את ה-F:



****

****

****

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | | סטטיסטי/מבחן |  | | אזור דחייה | |
| 1. מבחן לטיב התאמה   Goodness  of Fit | | הנתונים שייכים להתפלגות מסוימת | | k קבוצות  n תצפיות  Ei: כמות צפויה לפי H0 בקבוצה i  Oi : כמות בפועל בקבוצה i  Ei5 | הנתונים אינם שייכים להתפלגות המשוערת | |  | |
| t מספר הערכים הנחוצים לחישוב  Ei  שהוסקו מהמדגם | |
|  | |
| 2.לוחות תלות הדדית  Contingency  Tables | | שני המשתנים האיכותיים הם בלתי תלויים | | : סכום של השורה i  : סכום של העמודה j n: סכום כולל | שני המשתנים הם תלויים | | r מספר שורות  c מספר עמודות | |
| 1. ניתוח שונות   Analysis of Variance | | n : מספר תצפיות בכל קבוצה  k : כמות הקבוצות | | סכום התצפיות בקבוצה i      סכום כל התצפיות      SSW=SST – SSB |  | |  | |
| 4. ניתוח שונות  מספר לא אחיד של תצפיות בכל קבוצה | |  | | N במקום nk  ni במקום n  ni =מספר תצפיות בקבוצה i |  | |  | |
|  |  | סטטיסטי/מבחן | | |  | אזור דחייה |
| 5. רגרסיה | b=0 |  | | | b≠0 |  |
| מתאם  correlation | ρ=0 |  | | | ρ≠0 | על פי הטבלה |

**דף עזר להבנת פלטיSPSS**

**נערך על ידי ד"ר אלישבע קופליק**

דוגמא: יצרן תמרוקים רוצה לבדוק אם תוחלת משקלן של השפופרות הממולאות במפעלו היא אמנם 103 גרם. התקבלו הנתונים הבאים (ניתן להניח התפלגות נורמלית): 96, 102, 98, 102, 104, 98, 100. מה ניתן להסיק, ברמת מובהקות 0.02?

מסקנה: לא דוחים את *H0*, התוחלת איננה שונה מ-103 בצורה מובהקת.



**Test value=103**: זה הערך המופיע בהשערת האפס.

**t=-2.806** : הסטטיסטי שחישבנו לעיל.

**df=6** מספר דרגות החופש הוא 6 (n-1).

**Sig.(2- tailed)=0.031** זהו ערך ה P-value עבור מבחן דו-צדדי. התוכנה SPSS תמיד עורכת מבחן דו צדדי, ולכן במקרה שנרצה לבצע מבחן חד צדדי, נצטרך לחלק ערך זה ב-2 כדי לקבל את מובהקות התוצאה למבחן חד צדדי.

**Mean Difference** : נותן את ההפרש בין הממוצע לתוחלת המשוערת על פי .

**95% Confidence Interval of the Difference:**  במקום לחשב רווח סמך עבור התוחלת, התוכנה SPSS מחשבת רווח סמך ברמת בטחון 0.95 עבור ההפרש שבין התוחלת המשוערת לממוצע המדגם כך: . ולכן בדוגמה שלנו יוצא כלומר .

כדי למצוא מתוך זה רווח סמך רגיל צריך לכל אחד מהקצוות להוסיף את הערך ואז להוסיף .

**הערה**: התוכנהSPSS תמיד משתמשת במבחן t (היות והתפלגות t שואפת להתפלגות נורמאלית כאשר n גדול).

דוגמא: חוקר רוצה לבדוק אם קיים הבדל מובהק בין הישגים של בנים לבין הישגים של בנות במבחן בסטטיסטיקה. ידוע שציוני המבחן מתפלגים נורמלית (מניחים שסטיות התקן שוות, ).

מסקנה: דוחים את *H0*, קיים הבדל מובהק בין המגדרים.

****

****

הסבר הפלט Group Statistics:

לכל מגדר ציינו את גודל המדגם = N, את הציון הממוצע Mean= ואת סטיית התקן =Std. Deviation

הסבר הפלט Independent Samples Test:

התוכנה SPSS מבצעת קודם כל מבחן F לבדוק אם השונויות שוות או לא

(Levene's Test for Equality of Variances)

כלומר מתבצעת בדיקת השערות:

ערך ה p-value=0.134 . לנו נתנו שהוא קטן מה p-value, ולכן לא דוחים את , כלומר השונויות שוות.

כעת אפשר לעבור למבחן לבדיקת השוויון בין התוחלות.

אנחנו מתייחסים לשורה הראשונה (כי השונויות שוות) (Equal variances assumed).

ולכן דוחים את וקיים הבדל בין המגדרים.

אחרת, אם סטיות התקן לא שוות, צריך להתייחס לשורה השניה

(Equal variances not assume ) אזי הנוסחאות שונות במקצת:

אצלינו בדוגמא יוצא

דרגות החופש מחושבות לפי הנוסחא: ואם מציבים את הנתונים שלנו בטבלה אכן מגיעים למספר 25.993.

דוגמא: לקבוצת נבדקים נמסר שאלות לשם הבהרת עמדותיהם בנושא מסוים. לאחר מכן הוקרן לפניהם סרט, שהציג את הנושא בצורה חיובית, ושוב נתבקשו הנבדקים למלא שאלון זה. האם הגביר הסרט את האהדה לנושא? נתוני הנבדקים מוצגים להלן (, ניתן להניח כי ההפרש מתפלג נורמלית).

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|  | 16 | 18 | 20 | 24 | 24 | 22 | 20 | 18 | 18 | 8 | 20 |
|  | 24 | 20 | 24 | 28 | 30 | 20 | 24 | 22 | 10 | 18 | 24 |

מסקנה: דוחים את *H0*, ההשפעה מובהקת.

****

****

הסבר הפלט:

**Mean** = הממוצע

**N**= גודל המדגם

**Std. Deviation**= סטיית התקן המחושבת מתוך המדגם )

**Std. Error Mean**=

מחשבים את ההפרש: d=before-after

בדיקת טיב התאמה:

דוגמא: הטלת קובייה 120 פעם. רוצים לבחון אם הקובייה מאוזנת.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 6 | 5 | 4 | 3 | 2 | 1 |  |
| 24 | 19 | 18 | 17 | 22 | 20 |  |
| 20 | 20 | 20 | 20 | 20 | 20 |  |
| 16 | 1 | 4 | 9 | 4 | 0 |  |
|  |  |  |  |  | 0 |  |







מסקנה: לא דוחים את ההשערה

****

Observed N = הנתונים מהמדגם

Expected N = הנתונים המחושבים על סמך ההתפלגות המשוערת

Residual=

****



Df=5 =(k-t) וכיון ש k=6 ו- t=1, היות ואנו נדרשים רק ל-n כדי לחשב את השכיחויות הצפויות.

הפלט מוסיף הערה: אין תאים שבהם קטן מ- 5. וערך הנמוך ביותר של הוא 20.0.

לוח תלות הדדית:

דוגמא: האם קיימת תלות מובהקת, ברמת מובהקות 0.05, בין דת ואזור מגורים בארצות הברית?

****

נתונה טבלה דו-מימדית של דת לעומת איזור מגורים. בכל קוביה מופיע הערך מהמדגם ומתחתיו הערך שמצפים לקבל אם אין תלות - . למשל: קטולים שגרים במזרח: ו .

יש גם סיכום כללי Total של כל שורה לבד וכל עמודה לבד.

ביצוע מבחן חי-בריבוע:

****

 מחושב לפי הנוסחא בחוברת.

N of valid cases: גודל המדגם ה"כשר".

בהערה a אומרים לנו שאין תאים עם קטן מ 5 וכי הערך המינימלי של ה הוא 125.20.

ניתוח שונות – Anova

דוגמא: כדי לבחון את התועלת של 4 דשנים שונים, השתמשו בכל אחד ב-5 חלקות, והתקבלו היבולים הבאים (בק"ג). האם יש הבדל מובהק בין הדשנים השונים, ברמת מובהקות 0.05?

****

הפלט הוא של היבול Yevul=

SSB=Between Groups=225.750

SSW=Within Groups=1616.800

SST=Total=1842.550

Df= דרגות החופש של כל שורה

Mean Square הוא הערך של ה SS חלקי דרגות החופש (למשל: )

הסטטיסטי שמחשבים:*.*

Sig.=p-value=0.541 לכן עבור אי אפשר לדחות את , כלומר, אין הבדל מובהק בין הדשנים השונים.

רגרסיה:

דוגמה:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| גובה: X | 160 | 160 | 120 | 200 |
| משקל: Y | 75 | 90 | 60 | 75 |

חישוב ידני נותן:



****

פלט של הקורולציה: מטריצה ריבועית שבודקת קשר בין כל 2 משתנים. בשורה הראשונה רשום הערך של r=Pearson Correlation . מתחתיו מופיע ערך ה p-value ומתחתיו גודל המדגם N.

שימו לב שהקורולציה של כל ערך עם עצמו הוא 1.

R(x=height,y=weight)=0.5

דוגמה: ציון במתמטיקה ובסטטיסטיקה

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y | 80 | 80 | 60 | 60 |
| X | 60 | 75 | 50 | 40 |

****

מודל הרגרסיה: R=0.87 (מקדם המתאם)

R square=0.757 (R בריבוע)

****

Sum of squares Regression=SSR=302.804

Sum of squares Residual=SSE=97.196

Sum of squares Total=SST=SSR+SSE=400

Mean Square = SS/df

באופן ידני:

****

זהו פלט המחשב את המקדמים של הרגרסיה

.=constant=32.150

.=0.673

ולכן מודל הרגרסיה:.

באמצעות מבחן t ניתן לבחון את מובהקותם של המדמים, כאשר . מתברר כי אינו מובהק (sig = 0.175), וכן אינו מובהק (sig = 0.13) ברמת מובהקות 0.05.

## Language-Of-Greek2.gif

## "יפיותו של יפת": הוא לשון יון, לשונו יפה משל כל בני יפת (רש"י מגילה ט ע"ב).

## Table of the Normal Distribution

|  |
| --- |
| normal01Probability Content  from   - to z     **Z 0.00 0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09**  **----+----------------------------------------------------------------------**  **0.0 | 0.5000 0.5040 0.5080 0.5120 0.5160 0.5199 0.5239 0.5279 0.5319 0.5359**  **0.1 | 0.5398 0.5438 0.5478 0.5517 0.5557 0.5596 0.5636 0.5675 0.5714 0.5753**  **0.2 | 0.5793 0.5832 0.5871 0.5910 0.5948 0.5987 0.6026 0.6064 0.6103 0.6141**  **0.3 | 0.6179 0.6217 0.6255 0.6293 0.6331 0.6368 0.6406 0.6443 0.6480 0.6517**  **0.4 | 0.6554 0.6591 0.6628 0.6664 0.6700 0.6736 0.6772 0.6808 0.6844 0.6879**  **0.5 | 0.6915 0.6950 0.6985 0.7019 0.7054 0.7088 0.7123 0.7157 0.7190 0.7224**  **0.6 | 0.7257 0.7291 0.7324 0.7357 0.7389 0.7422 0.7454 0.7486 0.7517 0.7549**  **0.7 | 0.7580 0.7611 0.7642 0.7673 0.7704 0.7734 0.7764 0.7794 0.7823 0.7852**  **0.8 | 0.7881 0.7910 0.7939 0.7967 0.7995 0.8023 0.8051 0.8078 0.8106 0.8133**  **0.9 | 0.8159 0.8186 0.8212 0.8238 0.8264 0.8289 0.8315 0.8340 0.8365 0.8389**  **1.0 | 0.8413 0.8438 0.8461 0.8485 0.8508 0.8531 0.8554 0.8577 0.8599 0.8621**  **1.1 | 0.8643 0.8665 0.8686 0.8708 0.8729 0.8749 0.8770 0.8790 0.8810 0.8830**  **1.2 | 0.8849 0.8869 0.8888 0.8907 0.8925 0.8944 0.8962 0.8980 0.8997 0.9015**  **1.3 | 0.9032 0.9049 0.9066 0.9082 0.9099 0.9115 0.9131 0.9147 0.9162 0.9177**  **1.4 | 0.9192 0.9207 0.9222 0.9236 0.9251 0.9265 0.9279 0.9292 0.9306 0.9319**  **1.5 | 0.9332 0.9345 0.9357 0.9370 0.9382 0.9394 0.9406 0.9418 0.9429 0.9441**  **1.6 | 0.9452 0.9463 0.9474 0.9484 0.9495 0.9505 0.9515 0.9525 0.9535 0.9545**  **1.7 | 0.9554 0.9564 0.9573 0.9582 0.9591 0.9599 0.9608 0.9616 0.9625 0.9633**  **1.8 | 0.9641 0.9649 0.9656 0.9664 0.9671 0.9678 0.9686 0.9693 0.9699 0.9706**  **1.9 | 0.9713 0.9719 0.9726 0.9732 0.9738 0.9744 0.9750 0.9756 0.9761 0.9767**  **2.0 | 0.9772 0.9778 0.9783 0.9788 0.9793 0.9798 0.9803 0.9808 0.9812 0.9817**  **2.1 | 0.9821 0.9826 0.9830 0.9834 0.9838 0.9842 0.9846 0.9850 0.9854 0.9857**  **2.2 | 0.9861 0.9864 0.9868 0.9871 0.9875 0.9878 0.9881 0.9884 0.9887 0.9890**  **2.3 | 0.9893 0.9896 0.9898 0.9901 0.9904 0.9906 0.9909 0.9911 0.9913 0.9916**  **2.4 | 0.9918 0.9920 0.9922 0.9925 0.9927 0.9929 0.9931 0.9932 0.9934 0.9936**  **2.5 | 0.9938 0.9940 0.9941 0.9943 0.9945 0.9946 0.9948 0.9949 0.9951 0.9952**  **2.6 | 0.9953 0.9955 0.9956 0.9957 0.9959 0.9960 0.9961 0.9962 0.9963 0.9964**  **2.7 | 0.9965 0.9966 0.9967 0.9968 0.9969 0.9970 0.9971 0.9972 0.9973 0.9974**  **2.8 | 0.9974 0.9975 0.9976 0.9977 0.9977 0.9978 0.9979 0.9979 0.9980 0.9981**  **2.9 | 0.9981 0.9982 0.9982 0.9983 0.9984 0.9984 0.9985 0.9985 0.9986 0.9986**  **3.0 | 0.9987 0.9987 0.9987 0.9988 0.9988 0.9989 0.9989 0.9989 0.9990 0.9990** |

Student's t Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **df\p** | **0.40** | **0.25** | **0.10** | **0.05** | **0.025** | **0.01** | **0.005** | **0.0005** |
| **1** | 0.324920 | 1.000000 | 3.077684 | 6.313752 | 12.70620 | 31.82052 | 63.65674 | 636.6192 |
| **2** | 0.288675 | 0.816497 | 1.885618 | 2.919986 | 4.30265 | 6.96456 | 9.92484 | 31.5991 |
| **3** | 0.276671 | 0.764892 | 1.637744 | 2.353363 | 3.18245 | 4.54070 | 5.84091 | 12.9240 |
| **4** | 0.270722 | 0.740697 | 1.533206 | 2.131847 | 2.77645 | 3.74695 | 4.60409 | 8.6103 |
| **5** | 0.267181 | 0.726687 | 1.475884 | 2.015048 | 2.57058 | 3.36493 | 4.03214 | 6.8688 |
| **6** | 0.264835 | 0.717558 | 1.439756 | 1.943180 | 2.44691 | 3.14267 | 3.70743 | 5.9588 |
| **7** | 0.263167 | 0.711142 | 1.414924 | 1.894579 | 2.36462 | 2.99795 | 3.49948 | 5.4079 |
| **8** | 0.261921 | 0.706387 | 1.396815 | 1.859548 | 2.30600 | 2.89646 | 3.35539 | 5.0413 |
| **9** | 0.260955 | 0.702722 | 1.383029 | 1.833113 | 2.26216 | 2.82144 | 3.24984 | 4.7809 |
| **10** | 0.260185 | 0.699812 | 1.372184 | 1.812461 | 2.22814 | 2.76377 | 3.16927 | 4.5869 |
| **11** | 0.259556 | 0.697445 | 1.363430 | 1.795885 | 2.20099 | 2.71808 | 3.10581 | 4.4370 |
| **12** | 0.259033 | 0.695483 | 1.356217 | 1.782288 | 2.17881 | 2.68100 | 3.05454 | 4.3178 |
| **13** | 0.258591 | 0.693829 | 1.350171 | 1.770933 | 2.16037 | 2.65031 | 3.01228 | 4.2208 |
| **14** | 0.258213 | 0.692417 | 1.345030 | 1.761310 | 2.14479 | 2.62449 | 2.97684 | 4.1405 |
| **15** | 0.257885 | 0.691197 | 1.340606 | 1.753050 | 2.13145 | 2.60248 | 2.94671 | 4.0728 |
| **16** | 0.257599 | 0.690132 | 1.336757 | 1.745884 | 2.11991 | 2.58349 | 2.92078 | 4.0150 |
| **17** | 0.257347 | 0.689195 | 1.333379 | 1.739607 | 2.10982 | 2.56693 | 2.89823 | 3.9651 |
| **18** | 0.257123 | 0.688364 | 1.330391 | 1.734064 | 2.10092 | 2.55238 | 2.87844 | 3.9216 |
| **19** | 0.256923 | 0.687621 | 1.327728 | 1.729133 | 2.09302 | 2.53948 | 2.86093 | 3.8834 |
| **20** | 0.256743 | 0.686954 | 1.325341 | 1.724718 | 2.08596 | 2.52798 | 2.84534 | 3.8495 |
| **21** | 0.256580 | 0.686352 | 1.323188 | 1.720743 | 2.07961 | 2.51765 | 2.83136 | 3.8193 |
| **22** | 0.256432 | 0.685805 | 1.321237 | 1.717144 | 2.07387 | 2.50832 | 2.81876 | 3.7921 |
| **23** | 0.256297 | 0.685306 | 1.319460 | 1.713872 | 2.06866 | 2.49987 | 2.80734 | 3.7676 |
| **24** | 0.256173 | 0.684850 | 1.317836 | 1.710882 | 2.06390 | 2.49216 | 2.79694 | 3.7454 |
| **25** | 0.256060 | 0.684430 | 1.316345 | 1.708141 | 2.05954 | 2.48511 | 2.78744 | 3.7251 |
| **26** | 0.255955 | 0.684043 | 1.314972 | 1.705618 | 2.05553 | 2.47863 | 2.77871 | 3.7066 |
| **27** | 0.255858 | 0.683685 | 1.313703 | 1.703288 | 2.05183 | 2.47266 | 2.77068 | 3.6896 |
| **28** | 0.255768 | 0.683353 | 1.312527 | 1.701131 | 2.04841 | 2.46714 | 2.76326 | 3.6739 |
| **29** | 0.255684 | 0.683044 | 1.311434 | 1.699127 | 2.04523 | 2.46202 | 2.75639 | 3.6594 |
| **30** | 0.255605 | 0.682756 | 1.310415 | 1.697261 | 2.04227 | 2.45726 | 2.75000 | 3.6460 |
| **Inf** | 0.253347 | 0.674490 | 1.281552 | 1.644854 | 1.95996 | 2.32635 | 2.57583 | 3.2905 |

# Table: Chi-Square Probabilities

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **df** | **0.995** | **0.99** | **0.975** | **0.95** | **0.90** | **0.10** | **0.05** | **0.025** | **0.01** | **0.005** |
| **1** | --- | --- | 0.001 | 0.004 | 0.016 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 |
| **2** | 0.010 | 0.020 | 0.051 | 0.103 | 0.211 | 4.605 | 5.991 | 7.378 | 9.210 | 10.597 |
| **3** | 0.072 | 0.115 | 0.216 | 0.352 | 0.584 | 6.251 | 7.815 | 9.348 | 11.345 | 12.838 |
| **4** | 0.207 | 0.297 | 0.484 | 0.711 | 1.064 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 |
| **5** | 0.412 | 0.554 | 0.831 | 1.145 | 1.610 | 9.236 | 11.070 | 12.833 | 15.086 | 16.750 |
| **6** | 0.676 | 0.872 | 1.237 | 1.635 | 2.204 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 |
| **7** | 0.989 | 1.239 | 1.690 | 2.167 | 2.833 | 12.017 | 14.067 | 16.013 | 18.475 | 20.278 |
| **8** | 1.344 | 1.646 | 2.180 | 2.733 | 3.490 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 |
| **9** | 1.735 | 2.088 | 2.700 | 3.325 | 4.168 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 |
| **10** | 2.156 | 2.558 | 3.247 | 3.940 | 4.865 | 15.987 | 18.307 | 20.483 | 23.209 | 25.188 |
| **11** | 2.603 | 3.053 | 3.816 | 4.575 | 5.578 | 17.275 | 19.675 | 21.920 | 24.725 | 26.757 |
| **12** | 3.074 | 3.571 | 4.404 | 5.226 | 6.304 | 18.549 | 21.026 | 23.337 | 26.217 | 28.300 |
| **13** | 3.565 | 4.107 | 5.009 | 5.892 | 7.042 | 19.812 | 22.362 | 24.736 | 27.688 | 29.819 |
| **14** | 4.075 | 4.660 | 5.629 | 6.571 | 7.790 | 21.064 | 23.685 | 26.119 | 29.141 | 31.319 |
| **15** | 4.601 | 5.229 | 6.262 | 7.261 | 8.547 | 22.307 | 24.996 | 27.488 | 30.578 | 32.801 |
| **16** | 5.142 | 5.812 | 6.908 | 7.962 | 9.312 | 23.542 | 26.296 | 28.845 | 32.000 | 34.267 |
| **17** | 5.697 | 6.408 | 7.564 | 8.672 | 10.085 | 24.769 | 27.587 | 30.191 | 33.409 | 35.718 |
| **18** | 6.265 | 7.015 | 8.231 | 9.390 | 10.865 | 25.989 | 28.869 | 31.526 | 34.805 | 37.156 |
| **19** | 6.844 | 7.633 | 8.907 | 10.117 | 11.651 | 27.204 | 30.144 | 32.852 | 36.191 | 38.582 |
| **20** | 7.434 | 8.260 | 9.591 | 10.851 | 12.443 | 28.412 | 31.410 | 34.170 | 37.566 | 39.997 |
| **21** | 8.034 | 8.897 | 10.283 | 11.591 | 13.240 | 29.615 | 32.671 | 35.479 | 38.932 | 41.401 |
| **22** | 8.643 | 9.542 | 10.982 | 12.338 | 14.041 | 30.813 | 33.924 | 36.781 | 40.289 | 42.796 |
| **23** | 9.260 | 10.196 | 11.689 | 13.091 | 14.848 | 32.007 | 35.172 | 38.076 | 41.638 | 44.181 |
| **24** | 9.886 | 10.856 | 12.401 | 13.848 | 15.659 | 33.196 | 36.415 | 39.364 | 42.980 | 45.559 |
| **25** | 10.520 | 11.524 | 13.120 | 14.611 | 16.473 | 34.382 | 37.652 | 40.646 | 44.314 | 46.928 |
| **26** | 11.160 | 12.198 | 13.844 | 15.379 | 17.292 | 35.563 | 38.885 | 41.923 | 45.642 | 48.290 |
| **27** | 11.808 | 12.879 | 14.573 | 16.151 | 18.114 | 36.741 | 40.113 | 43.195 | 46.963 | 49.645 |
| **28** | 12.461 | 13.565 | 15.308 | 16.928 | 18.939 | 37.916 | 41.337 | 44.461 | 48.278 | 50.993 |
| **29** | 13.121 | 14.256 | 16.047 | 17.708 | 19.768 | 39.087 | 42.557 | 45.722 | 49.588 | 52.336 |
| **30** | 13.787 | 14.953 | 16.791 | 18.493 | 20.599 | 40.256 | 43.773 | 46.979 | 50.892 | 53.672 |

**Table of the Upper Percentage Points of the F(u,v) distribution**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| v | u | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 |
| 1 | .05 .01 | 161 4052 | 200 4999 | 216 5403 | 225 5625 | 230 5764 | 234 5859 | 237 5928 | 239 5981 | 241 6022 | 242 6056 | 243 6082 | 244 6106 | 246 6157 |
| 2 | .05 .01 | 18.51 98.49 | 19.00 99.01 | 19.16 99.17 | 19.25 99.25 | 19.39 99.30 | 19.33 99.33 | 19.36 99.34 | 19.37 99.36 | 19.38 99.38 | 19.39 99.40 | 19.4 99.4 | 19.4 99.4 | 19.4 99.4 |
| 3 | .05 .01 | 10.13 34.12 | 9.55 30.81 | 9.28 29.46 | 9.12 28.71 | 9.01 28.24 | 8.94 27.91 | 8.88 27.67 | 8.84 27.49 | 8.81 27.34 | 8.78 27.23 | 8.76 27.13 | 8.74 27.05 | 8.70 26.87 |
| 4 | .05 .01 | 7.71 21.20 | 6.94 18.00 | 6.59 16.69 | 6.39 15.98 | 6.26 15.52 | 6.16 15.21 | 6.09 14.98 | 6.04 14.80 | 6.00 14.66 | 5.96 14.54 | 5.93 14.95 | 5.91 14.37 | 5.86 14.2 |
| 5 | .05 .01 | 6.61 16.26 | 5.79 13.27 | 5.41 12.06 | 5.19 11.39 | 5.05 10.97 | 4.95 10.67 | 4.88 10.45 | 4.82 10.27 | 4.78 10.15 | 4.74 10.05 | 4.70 9.96 | 4.68 9.89 | 4.62 9.72 |
| 6 | .05 .01 | 5.99 13.74 | 5.14 10.92 | 4.76 9.78 | 4.53 9.15 | 4.39 8.75 | 4.28 8.47 | 4.21 8.26 | 4.15 8.10 | 4.10 7.98 | 4.06 7.87 | 4.03 7.79 | 4.00 7.72 | 3.94 7.56 |
| 7 | .05 .01 | 5.59 12.25 | 4.74 9.55 | 4.35 8.45 | 4.12 7.85 | 3.97 7.46 | 3.87 7.19 | 3.79 7.00 | 3.73 6.84 | 3.68 6.71 | 3.63 6.62 | 3.60 6.54 | 3.57 6.47 | 3.51 6.31 |
| 8 | .05 .01 | 5.32 11.26 | 4.46 8.65 | 4.07 7.59 | 3.84 7.01 | 3.69 6.63 | 3.58 6.37 | 3.50 6.19 | 3.44 6.03 | 3.39 5.91 | 3.34 5.82 | 3.31 5.74 | 3.28 5.67 | 3.22 5.52 |
| 9 | .05 .01 | 5.12 10.56 | 4.26 8.02 | 3.86 6.99 | 3.63 6.42 | 3.48 6.06 | 3.37 5.80 | 3.29 5.61 | 3.23 5.47 | 3.18 5.35 | 3.14 5.26 | 3.10 5.18 | 3.07 5.11 | 3.01 4.96 |
| 10 | .05 .01 | 4.96 10.04 | 4.10 7.56 | 3.71 6.55 | 3.48 5.99 | 3.33 5.64 | 3.22 5.39 | 3.14 5.21 | 3.07 5.06 | 3.02 4.95 | 2.97 4.85 | 2.94 4.78 | 2.91 4.71 | 2.85 4.56 |
| 20 | .05 .01 | 4.30 7.95 | 3.44 5.72 | 3.05 4.82 | 2.82 4.31 | 2.66 3.99 | 2.55 3.76 | 2.46 3.59 | 2.40 3.45 | 2.34 3.35 | 2.30 3.26 | 2.31 3.30 | 2.28 3.23 | 2.20 3.09 |
| 30 | .05 .01 | 4.17 7.56 | 3.32 5.39 | 2.92 4.51 | 2.69 4.02 | 2.53 3.70 | 2.42 3.47 | 2.33 3.30 | 2.27 3.17 | 2.21 3.07 | 2.16 2.98 | 2.12 2.90 | 2.09 2.84 | 2.01 2.70 |
| 40 | .05 .01 | 4.08 7.31 | 3.23 5.18 | 2.84 4.31 | 2.61 3.83 | 2.45 3.51 | 2.34 3.29 | 2.25 3.12 | 2.18 2.99 | 2.12 2.89 | 2.08 2.80 | 2.04 2.73 | 2.00 2.66 | 1.92 2.52 |
| 50 | .05 .01 | 4.03 7.17 | 3.18 5.06 | 2.79 4.20 | 2.56 3.72 | 2.40 3.41 | 2.29 3.19 | 2.20 3.02 | 2.13 2.89 | 2.07 2.78 | 2.03 2.70 | 1.98 2.62 | 1.95 2.56 | 1.87 2.42 |
| 60 | .05 .01 | 4.00 7.08 | 3.15 4.98 | 2.76 4.13 | 2.53 3.65 | 2.37 3.34 | 2.25 3.12 | 2.17 2.95 | 2.10 2.82 | 2.04 2.72 | 1.99 2.63 | 1.95 2.56 | 1.92 2.50 | 1.84 2.35 |
| 70 | .05 .01 | 3.98 7.01 | 3.13 4.92 | 2.74 4.08 | 2.50 3.60 | 2.35 3.29 | 2.23 3.07 | 2.14 2.91 | 2.07 2.77 | 2.01 2.67 | 1.97 2.59 | 1.93 2.51 | 1.89 2.45 | 1.82 2.32 |
| 80 | .05 .01 | 3.96 6.96 | 3.11 4.88 | 2.72 4.04 | 2.48 3.56 | 2.33 3.25 | 2.21 3.04 | 2.12 2.87 | 2.05 2.74 | 1.99 2.64 | 1.95 2.55 | 1.91 2.48 | 1.88 2.41 | 1.80 2.28 |
| 100 | .05 .01 | 3.94 6.90 | 3.09 4.82 | 2.70 3.98 | 2.46 3.51 | 2.31 3.21 | 2.19 2.99 | 2.10 2.82 | 2.03 2.69 | 1.97 2.59 | 1.93 2.50 | 1.88 2.43 | 1.85 2.36 | 1.77 2.22 |
| 200 | .05 .01 | 3.89 6.76 | 3.04 4.71 | 2.65 3.88 | 2.42 3.41 | 2.26 3.11 | 2.14 2.89 | 2.06 2.73 | 1.98 2.60 | 1.93 2.50 | 1.88 2.41 | 1.83 2.34 | 1.80 2.28 | 1.72 2.13 |
| 400 | .05 .01 | 3.86 6.70 | 3.02 4.66 | 2.62 3.83 | 2.39 3.36 | 2.23 3.06 | 2.12 2.85 | 2.03 2.69 | 1.96 2.55 | 1.90 2.46 | 1.85 2.37 | 1.81 2.29 | 1.78 2.23 | 1.70 2.08 |
| 1000 | .05 .01 | 3.85 6.66 | 3.00 4.63 | 2.61 3.80 | 2.38 3.34 | 2.22 3.04 | 2.11 2.82 | 2.02 2.66 | 1.95 2.53 | 1.89 2.43 | 1.84 2.34 | 1.80 2.26 | 1.76 2.20 | 1.68 2.06 |
| Inf | .05 .01 | 3.85 6.64 | 2.99 4.60 | 2.60 3.78 | 2.37 3.32 | 2.21 3.02 | 2.09 2.80 | 2.01 2.64 | 1.94 2.51 | 1.88 2.41 | 1.83 2.32 | 1.79 2.24 | 1.75 2.18 | 1.66 2.03 |

Critical Values of the   
Pearson Correlation Coefficient

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| df = *n* -2 |  | | | |
| Level of Significance ()  for Two-Tailed Test | **.10** | **.05** | **.02** | **.01** |
| Df |  |  |  |  |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30  35  40  45  50  60  70  80  90  100 | .988  .900  .805  .729  .669  .622  .582  .549  .521  .497  .476  .458  .441  .426  .412  .400  .389  .378  .369  .360  .352  .344  .337  .330  .323  .317  .311  .306  .301  .296  .275  .257  .243  .231  .211  .195  .183  .173  .164 | .997  .950  .878  .811  .754  .707  .666  .632  .602  .576  .553  .532  .514  .497  .482  .468  .456  .444  .433  .423  .413  .404  .396  .388  .381  .374  .367  .361  .355  .349  .325  .304  .288  .273  .250  .232  .217  .205  .195 | .9995  .980  .934  .882  .833  .789  .750  .716  .685  .658  .634  .612  .592  .574  .558  .542  .528  .516  .503  .492  .482  .472  .462  .453  .445  .437  .430  .423  .416  .409  .381  .358  .338  .322  .295  .274  .256  .242  .230 | .9999  .990  .959  .917  .874  .834  .798  .765  .735  .708  .684  .661  .641  .623  .606  .590  .575  .561  .549  .537  .526  .515  .505  .496  .487  .479  .471  .463  .456  .449  .418  .393  .372  .354  .325  .303  .283  .267  .254 |

**הוראות הפעלת הפונקציות הסטטיסטיות שבמחשבון (*fx-82MS* CASIO)**

1. **חישוב מדדים שימושיים**

**שלב 1:**

כניסה למערכת הסטטיסטיקה: mode (2\SD)

וודא שעל מסך המחשבון מופיע למעלה הסימון SD.

**שלב 2:**

איפוס זיכרון הנתונים: shift-mode (1\Scl). ואז נקבל על מסך המחשבון את ההודעה Stat Clear . לוחצים אז על הסימן =, ואחר כך על AC במטרה לנקות את המסך.

**שלב 3:**

הזנת נתונים:

נתחיל להזין נתונים כאשר אחרי כל ערך נקיש 'M+', ובסיום הזנת הנתונים יש לנעול את הזיכרון ע"י המקש 'AC'.

ברירת המחדל היא שהשכיחות של כל נתון הוא 1. אם השכיחות גדולה מ-1, אחרי הקלדת המספר יש ללחוץ על shift-',' - כך שיופיע הסימן ';' - ואחר כך להקליד את השכיחות וללחוץ על 'M+'.

**שלב 4:**

על ידי לחיצת shift S-SUM (היינו shift 1), ניתן לקבל שלוש נוסחאות:

1: . הכוונה לסכום ריבועי הנתונים שנכנסו, היינו 

2: . הכוונה לסכום הנתונים שנכנסו, היינו .

3: n: הכוונה לכמות הנתונים שנכנסו.

על ידי לחיצת shift S-VAR (היינו shift 2), ניתן לקבל שלוש נוסחאות:

1: : ממוצע הנתונים שנכנסו.

2: : סטיית התקן בשימוש בסטטיסטיקה תיאורית, היינו: 

3: : סטיית התקן בשימוש בסטטיסטיקה הסקתית, היינו: 

1. **מקדם מתאם ורגרסיה**

**שלב 1:**

כניסה למערכת הרגרסיה: mode (3\REG), ואחר כך (1\Lin) . הכוונה לרגרסיה לינארית פשוטה.

וודא שעל מסך המחשבון מופיע למעלה הסימון REG.

**שלב 2:**

איפוס זיכרון הנתונים: shift-mode (1\Scl). ואז נקבל על מסך המחשבון את ההודעה Stat Clear . לוחצים אז על הסימן =, ואחר כך על AC במטרה לנקות את המסך.

**שלב 3:**

הזנת נתונים:

נתחיל להזין זוגות של נתונים (קודם כל x ואחר כך y) המופרדים על ידי פסיק, כאשר אחרי כל זוג נקיש 'M+', ובסיום הזנת הנתונים יש לנעול את הזיכרון ע"י המקש 'AC'.

ברירת המחדל היא שהשכיחות של כל זוג נתונים הוא 1. אם השכיחות גדולה מ-1, אחרי הקלדת זוג הנתונים יש ללחוץ על shift-',' - כך שיופיע הסימן ';' - ואחר כך להקליד את השכיחות וללחוץ על 'M+'.

**שלב 4:**

על ידי לחיצת shift S-VAR (היינו shift 2), ניתן לקבל תשע נוסחאות, כדלהלן:

1: : ממוצע הנתונים שנכנסו (משתנה מסביר).

2: : סטיית התקן של המשתנה המסביר, בשימוש בסטטיסטיקה תיאורית, היינו: 

3: : סטיית התקן של המשתנה המסביר, בשימוש בסטטיסטיקה הסקתית, היינו: 

4: : ממוצע הנתונים שנכנסו (משתנה מוסבר).

5: : סטיית התקן של המשתנה המוסבר, בשימוש בסטטיסטיקה תיאורית, היינו: 

6: : סטיית התקן של המשתנה המוסבר, בשימוש בסטטיסטיקה הסקתית, היינו: 

7: A: הערך של הקבוע a של נוסחת הרגרסיה.

8: B: הערך של המקדם b של נוסחת הרגרסיה.

9: r: מקדם המתאם של פירסון בין x ל-y.

**הוראות הפעלת הפונקציות הסטטיסטיות שבמחשבון (*fx-82ES* CASIO)**

1. **חישוב מדדים שימושיים**

**שלב 1:**

איפוס זיכרון הנתונים: shift-clr(9) ואז לוחצים על 3:All. ואחר כך: [=]:Yes, ואחר כך על AC.

**שלב 2:**

הוספת טור שכיחויות: shift-setup. ואז, לוחצים על replay ובוחרים 3:STAT. המחשבון שואל: Frequency?, ובוחרים 1:ON.

**שלב 3:**

הזנת נתונים:

לוחצים על MODE, ואחר כך על 2:STAT, ואחר כך על 1:1-VAR, ואז מופיעה טבלת שכיחויות שיש למלא. מכניסים כל נתון ואחר כך לוחצים על =. הנתון ייכנס בטבלה עם שכיחות 1 (זו ברירת המחדל). אם השכיחות שונה מ-1, ניתן למקם את הסמן על ידי הכפתור REPLAY במקום הנכון, ולהכניס את השכיחות על ידי =. בסוף התהליך יש ללחוץ על AC כדי לנקות את המסך.

**שלב 4:**

על ידי לחיצת shift- STAT (היינו shift-1), ואחר כך על 4:Sum, ניתן לקבל שתי נוסחאות:

1: . הכוונה לסכום ריבועי הנתונים שנכנסו, היינו 

2: . הכוונה לסכום הנתונים שנכנסו, היינו .

על ידי לחיצת shift- STAT (היינו shift-1), ואחר כך על 5:Var, ניתן לקבל ארבע נוסחאות:

1: n: הכוונה לכמות הנתונים שנכנסו.

2: : ממוצע הנתונים שנכנסו.

3: : סטיית התקן בשימוש בסטטיסטיקה תיאורית, היינו: 

4: : סטיית התקן בשימוש בסטטיסטיקה הסקתית, היינו: 

1. **מקדם מתאם ורגרסיה**

**שלב 1:**

איפוס זיכרון הנתונים: shift-clr(9) ואז לוחצים על 3:All. ואחר כך: [=]:Yes, ואחר כך על AC.

**שלב 2:**

הוספת טור שכיחויות: shift-setup. ואז, לוחצים על replay ובוחרים 3:STAT. המחשבון שואל: Frequency?, ובוחרים 1:ON.

**שלב 3:**

הזנת נתונים:

לוחצים על MODE, ואחר כך על 2:STAT, ואחר כך על 2:A+BX, ואז מופיעה טבלת שכיחויות, לכל זוג X,Y, שיש למלא. מכניסים כל נתון ואחר כך לוחצים על =. הנתון ייכנס בטבלה עם שכיחות 1 (זו ברירת המחדל). אם השכיחות שונה מ-1, ניתן למקם את הסמן על ידי הכפתור REPLAY במקום הנכון, ולהכניס את השכיחות על ידי =. בסוף התהליך יש ללחוץ על AC כדי לנקות את המסך.

**שלב 4:**

על ידי לחיצתshift-STAT (היינו shift-1), ואחר כך על 7:Reg, ניתן לקבל חמש נוסחאות, שרק שלוש מהן רלוונטיות, כדלהלן:

1: A: הערך של הקבוע a של נוסחת הרגרסיה.

2: B: הערך של המקדם b של נוסחת הרגרסיה.

3: r: מקדם המתאם של פירסון בין x ל-y.

הערה: על ידי לחיצת 4:Sum או 5:Var ניתן לקבל מדדים שונים, בדומה למה שמוזכר בסעיף הקודם.