

079

החוג לתקשורת

1/10

אוניברסיטת חיפה

13

מבחן סיום בקורס
"סטטיסטיקה ממוחשבת"

סמסטר ב'

10

19

מרצה: יעל קורש
מתרגלים: איריס זמיר ורוני טוטיאן

תשנ"ח - אוגוסט, 1998

מועד ב'

חומר עזר מותר: ספרים, מחברות ומחשב כיס

משך הבחינה: שעתיים

יש לענות על השאלות בטופס הבחינה
במידת הצורך ניתן לתלוש את הדף האחרון, אך יש להגישו עם טופס הבחינה

בהצלחה!



כסטודנט/ית בחוג לתקשורת אשר סיים/ה לפני כחודש את הקורס "סטטיסטיקה ממוחשבת", הופנית לעבוד כאסיסטנט לסטטיסטיקאי הראשי פרופ' קשקשת מהמכון לחקר האנדרלמוסיה בתקשורת. כמי שהתמחה/תה בקורס בנושא יסודות הסטטיסטיקה התיאורית וההסקתית ועיבוד נתונים באמצעות המחשב (בתוכנה הסטטיסטית SPSS), נתבקשת לעזור ולייעץ לפרופ' קשקשת בעיבוד הנתונים של סקר בנושא: הרגלי הצפייה בטלוויזיה של האוכלוסייה בישראל בזמן שידור משחקי המונדיאל, צרפת 1998.

פרופ' קשקשת נתן לך קובץ נתונים שהוא ערך לאחר סקר טלפוני שערך במדגם מייצג של משקי בית בישראל. קובץ הנתונים כלל 1000 ראשי משק בית. פרופ' קשקשת צרף דף המפרט את המשתנים בקובץ (ראה פירוט בדף האחרון).

בשלב ראשון עליך להראות לפרופ' קשקשת כי הינך מבין/ה את החומר שנלמד בשיעורי ה"סטטיסטיקה הממוחשבת" בנוגע לסטטיסטיקה התיאורית וסטטיסטיקה הסקתית (60 נקודות עבור חלק זה). בשלב שני תצטרך/י לתת הסברים מפורטים ל-4 פלטים ממוחשבים (10 נקודות על הסבר מלא של כל פלט. סה"כ 40 נקודות עבור חלק זה).

לאחר קבלת קובץ הנתונים, פרופ' קשקשת אמר לך כי השערת המחקר המרכזית שלו הינה: **לגברים בעלי רמת השכלה תיכונית ומעלה הכנסה ממוצעת גבוהה יותר מנשים בעלות רמת השכלה דומה**

שאלתו הראשונה לגבי הממצאים היתה: (6 נקודות על תשובה מלאה)

(1.1) מהו המשתנה התלוי בהשערה זו?

(1.2) מהו המשתנה או המשתנים הבלתי תלויים בהשערה? נמק/י

לאחר שהכנת קובץ SPSS לקריאת ועיבוד הנתונים שנמסרו לך, והרצת כמה פרוצדורות לסטטיסטיקה תיאורית, חזרת לפרופ' קשקשת.

(2) שאלתו השניה התייחסה לגיל הנחקרים במדגם. לצורך כך, יצרת משתנה חדש בשם גיל. כיצד? (2 נקודות)

(7) ביחס לממוצע ההכנסה במדגם, ובהנחה כי התפלגות ההכנסה הינה נורמלית, אם נחקר דיווח כי סיכויי ההכנסה שלו הם 95%, מיקומו בהתפלגות הכללית של ההכנסות הוא: יש לסמן תשובה אחת בלבד (4 נקודות)

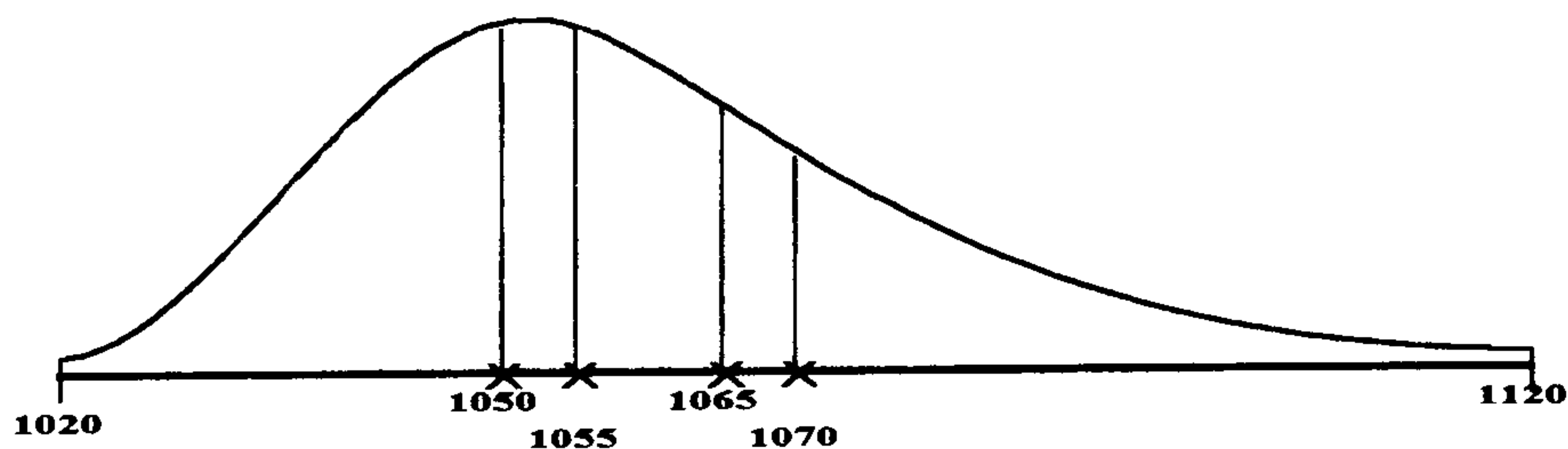
- קרוב יחסית לממוצע ההתפלגות
- בערך סטיית תקן אחת פחות מהממוצע
- בערך שתי סטיות תקן מעל הממוצע
- בערך שתי סטיות תקן מתחת לממוצע

שאלתנו הבאה התייחסה לממוצע ההכנסה בין קבוצות שונות במדגם: (6 נקודות על תשובה מלאה)
 (8) נתון כי ממוצע ההכנסה של 250 גברים בעלי השכלה יסודית בלבד הינה 2500 ש"ח, ממוצע ההכנסה של 320 גברים בעלי רמת השכלה תיכונית ומעלה הינה 3580 ש"ח, ממוצע ההכנסה של 150 נשים בעלות השכלה יסודית הינה 1500 ש"ח, ושל שאר הנשים במדגם (N=270) 2900 ש"ח. להזכירך, סה"כ הנחקרים במדגם הינו 1000 (N=1000).

א. האם תוכלי לחשב את ממוצע שעות הצפיה של כלל הנחקרים במדגם? אם כן, חשבי; אם לא – נמקי מדוע

ב. במידה ופרופ' קשקשת היה נותן לך חציון במקום ממוצע שעות צפיה במשחקי המונדיאל של כל קבוצה, האם היית יכולה לחשב את החציון הכללי? אם כן – חשבי; אם לא – נמקי מדוע לא

(9) בהתפלגות הבאה סומנו מקומותיהם של ששה מדדי פיזור וערכים מרכזיים המתייחסים להכנסה: ציניקי ליד חמישה מהם לאיזה מדד הוא מתאים. הסברי (5 נקודות)



אח"כ הכנת טבלת שכיחויות של התפלגות הנחקרים במדגם לפי קבוצות גיל:

שכיחות (f)	קבוצת גיל
150	0-10
260	11-20
390	21-40
200	41-80

(3) האם התפלגות הגילאים במדגם הינה נורמלית? נמק/י (3 נקודות)

(4) שרטט/י דיאגרמה לתיאור גרפי של התפלגות הנחקרים במדגם לפי קבוצות גיל (4 נקודות).

(5) לפניך נתונים לגבי שעות צפיה ביום של 11 נחקרים מקריים מהמדגם:

8 4 2 7 12 17 9 14 11 14 14

התברר כי חלה טעות ברישום מספר שעות הצפיה של הנחקר בעל 2 שעות צפיה, והוא למעשה צופה 5 שעות ביום. הסבר/י, ללא חישובים, כיצד ישפיע התיקון על הממוצע, השכיח, החציון וסטיית התקן של שעות הצפיה של קבוצת נחקרים זו (4 נקודות)

(6) פרופ' קשקשת ביקש שתציין/י לגבי ארבעת ההגדים הבאים האם הם נכונים או לא נכונים: (4 נקודות)

א. הציונים במבחן פסיכומטרי הינם ערכים של משתנה הנמדד בסולם אורדינלי

נכון / לא נכון

ב. חשיבותן של דרגות חופש לא קשור כלל בחישוב מובהקות הערך הסטטיסטי והשוואתו לערך קריטי בלוחות של מבחני המובהקות

נכון / לא נכון

ג. השונות היא סכום ריבועי הסטיות כאשר הסטיות נמדדות מהחציון

נכון / לא נכון

ד. אם H_0 לא נדחתה עבור רמת מובהקות של 0.05, היא בודאי לא תדחה עבור רמת מובהקות של 0.01

נכון / לא נכון

(10) התיאורים בלוחות השכיחות, בגרפים, הערכים המרכזיים ומדדי פיזור שהצגת, לא אפשרו לפרופ' קשקשת לאשש או להפריך את השערתו המרכזית. מכיוון שכך, הוא ביקש ממך להציג לו כמה ממצאים שיאפשרו לו לאשש או להפריך את השערתו. להלן הממצאים שפרופ' קשקשת ביקש ממך להציג. עליך לענות על כל שאלותיו, לספק נימוק לתשובתך, לציין את המדד או המבחן המתאים ולפרט לפחות שלושה ממצאים שיוצגו בשלב מאוחר יותר בפלט המחשב שתציג לו. אין לכלול כאחד משלושת הממצאים את גודל המדגם, מדדי פיזור או ערכים מרכזיים. עליך להתייחס אך ורק למשתנים המופיעים בהשערה המרכזית:

לגברים בעלי רמת השכלה תיכונית ומעלה הכנסה ממוצעת גבוהה יותר מנשים בעלות רמת השכלה דומה

א. האם קיים קשר בין המשתנה התלוי לבין כל אחד מהמשתנים הבלתי-תלויים בנפרד? (4 נקודות)

ב. במטרה לערוך טבלת שכיחויות זו-משתנית (בדומה לטבלת צילוח) למין ורמת השכלה: (סמך/י רק תשובה אחד - 2 נקודות)

1. די ברשימת התצפיות (ההתפלגות) של המשתנה "מין" ורשימת התצפיות (ההתפלגות) של המשתנה "רמת השכלה"

2. יש צורך ברשימת זוגות של תצפיות (ההתפלגות המשותפת) על "מין" ו"רמת השכלה"

ג. מהי מידת הניבוי של המשתנה התלוי על-ידי כל המשתנים הבלתי-תלויים ביחד? (4 נקודות)

ד. מהם התנאים בהם צריך לעמוד המחקר של פרופ' קשקשת על-מנת שיוכל להסיק מסקנות תקפות ומהימנות מהמדגם לכלל אוכלוסיית ישראל? ציין/י את התנאים ואת המדד או המבחן המתאים לכל תנאי (6 נקודות).

024

ה. האם קיימים הבדלים בין שתי הקבוצות הללו מבחינת ממוצע המשתנה התלוי והאם ההבדלים הללו מובהקים מבחינה סטטיסטית? ציין/י את המדד או המבחן המתאים, שתי הנחות יסוד לצורך השימוש במדד או במבחן ולפחות שלושה ממצאים שבכוונתך להציג מתוך פלט המחשב (6 נקודות).

עד לשלב זה, פרופי קשקשת מאד מרוצה מעבודתך. מכיוון שהצלחת להרשים אותו בידע וביצירתיות בחלק הקודם, הוא הטיל עליך משימה אחראית עוד יותר: ניתוח פלטים. על-מנת להראות לו שאכן את/ה ראוי/ה לאמונו, בחלק זה עליך להציג בפני פרופי קשקשת את הממצאים העיקריים המוצגים בכל פלט. זכור/י כי את/ה המומחה בסטטיסטיקה, ושעליך לרשום את הממצאים בצורה ברורה ומלאה, כך שהפרופסור יוכל להבין אותם....

076

פלט מספר 2

צייני את התפריטים העיקריים שבאמצעותם התקבל הפלט. כמו כן, צייני מהו המשתנה התלוי ומהם המשתנים הבלתי-תלויים, ולפחות 5 ממצאים עיקריים המוצגים בפלט - כולל התייחסות לכל אחד מהמשתנים הבלתי-תלויים וההבדלים ביניהם.

t-tests for Independent Samples of SEX

Variable	Number of Cases	Mean	SD	SE of Mean
ED				
SEX 1	484	12.1880	1.474	.067
SEX 2	516	14.8450	1.437	.063

Mean Difference = 2.657

Levene's Test for Equality of Variances: F= .881 P= .001

t-test for Equality of Means				95%	
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	3.73	998	.001	.042	(.162, .524)
Unequal	3.72	990.04	.001	.042	(.162, .524)

Variable	Number of Cases	Mean	SD	SE of Mean
INC				
SEX 1	484	2885.9050	1058.205	48.100
SEX 2	516	2871.2946	1053.772	46.390

Mean Difference = 14.6104

Levene's Test for Equality of Variances: F= .123 P= .726

t-test for Equality of Means				95%	
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	.22	998	.827	66.816	(-116.506, 145.727)
Unequal	.22	993.37	.827	66.825	(-116.525, 145.746)

027

פלט מספר 3

תארו והסבירו את כל הקשרים המופיעים בפלט להלן.

- - PARTIAL CORRELATION COEFFICIENTS - -

Zero Order Partial

	ED	GIL	REL	SEX
ED	1.0000 (0) P= .	.0298 (959) P= .001	.0211 (959) P= .257	-.1281 (959) P= .001
GIL	.0298 (959) P= .001	1.0000 (0) P= .	-.0285 (959) P= .189	-.0276 (959) P= .196
REL	.0211 (959) P= .257	-.0285 (959) P= .189	1.0000 (0) P= .	-.0082 (959) P= .399
SEX	-.1281 (959) P= .001	-.0276 (959) P= .196	-.0082 (959) P= .399	1.0000 (0) P= .

(Coefficient / (D.F.) / 1-tailed Significance)

" . " is printed if a coefficient cannot be computed

Controlling for..	REL	SEX
	ED	GIL
ED	1.0000 (0) P= .	.5271 (957) P= .0001
GIL	.5271 (957) P= .0001	1.0000 (0) P= .

(Coefficient / (D.F.) / 1-tailed Significance)

" . " is printed if a coefficient cannot be computed

028

- 9 -

פלט מספר 4

בפלט זה עליך לכתוב את התפריטים ב- SPSS באמצעותם התקבל הפלט, לציון את המשתנים ולהתייחס למרכיבים הבאים: טווח, ערך מינימלי, ערך מקסימלי, ממוצע, סטיית-תקן, טעות תקן, שונות.

Variable TVTIME

Mean	13.693	S.E. Mean	.118
Std Dev	3.727	Variance	13.889
Range	36.000	Minimum	8
Maximum	44	Sum	13693.000

Valid observations - 1000 Missing observations - 0

Variable ED

Mean	15.011	S.E. Mean	.046
Std Dev	1.465	Variance	2.145
Range	8.000	Minimum	12
Maximum	20	Sum	15011.000

Valid observations - 1000 Missing observations - 0

Variable GIL

Mean	43.159	S.E. Mean	.423
Std Dev	13.377	Variance	178.939
Range	44.000	Minimum	20.00
Maximum	64.00	Sum	43159.000

Valid observations - 1000 Missing observations - 0

Number of valid observations (listwise) = 1000.00

Variable INC

Mean	2878.366	S.E. Mean	33.375
Std Dev	1055.416	Variance	1113903.502
Range	6396.000	Minimum	2345
Maximum	8741	Sum	2878366.000

Valid observations - 1000 Missing observations - 0

פירוט המשתנים בקובץ הנתונים

1. שנת לידה - YBIRTH □ □ 19
2. מין - SEX .1 זכר
.2 נקבה
3. מספר ילדים - NCHILD □ □
4. הכנסה - INC □ □ □ □ □ ש"ח
5. מצב משפחתי - MSTAT
S - רווק
M - נשוי
W - אלמן
D - גרוש
6. השכלה - ED □ □ מספר שנות לימוד
7. רמת השכלה - SCHOOL
1. יסודית - חטיבת ביניים
2. תיכונית ללא תעודת בגרות
3. תיכונית עם תעודת בגרות
4. על-תיכונית
8. מוצא - MOZA
1. ילידי אסיה-אפריקה
2. ילידי אירופה-אמריקה
3. ילידי ישראל - דור ראשון ושני
9. רמת דתיות - REL
1. חילוני
2. מסורתי
3. דתי / חרדי
10. מספר שעות צפייה בטלוויזיה בשבוע של המונדיאל - TVTIME □ □ נע בין 1 ל- 25
11. סוג תוכנית בטלוויזיה - SHOW
1. סרט
2. אופרת סבון
3. תוכנית בידור
4. משחקי המונדיאל
5. חדשות
6. אחר
12. ערוץ צפייה - AROTZ
1. ערוץ 1
2. ערוץ 2
3. ערוץ מקומי
4. ערוץ הילדים
5. ערוץ הספורט
6. ערוץ הסרטים
7. ערוצים אחרים בעברית
8. ערוצים אחרים בשפה האנגלית
9. ערוצים בשפה הערבית
10. ערוצים בשפה אחרת