

שיטות סטטיסטיות לחשבונאים

מרצה: ד"ר דוד סעד

מתרגל: מר ליאור מיסטריל

סמסטר א' תשע"א

מועד הבחינה:	יום שישי 4.2.2011, שעה 08:30
משך הבחינה:	שלוש שעות
מס' עמודים:	4 עמודים
חומר עזר:	מותר השימוש בכל חומר עזר.

- נא לענות על כל השאלות (משקלן של השאלות זהה).
- נא לכתוב את כל החישובים במחברת הבחינה עצמה.

0-30

בהצלחה !!

הוראות לצורך סריקת מחברת הבחינה:

- נא להמנע מכתובה בעט ירוק או בעפרון.
- נא לא לכתוב בתחום השוליים.

בתום הבחינה על התלמיד להקפיד ולמסור למשגיחה באופן אישי את גיליון ומחברות הבחינה, ולהמתין עד אשר תסמן המשגיחה את המסירה.

1. א. שלטונות הצבא טוענים כי השתלבות המתגייסים ליחידות השונות תלויה מאד בתיכונים בהם למדו.

להלן נתוני מדגם שנעשה לאחרונה :

	חיל אויר וים	יחידות שירות	יחידות חי"ר	שריון וארטילריה
תיכונים בערים	35	15	29	11
המגזר הדרוזי	14	9	40	7
ישיבות תיכוניות	5	20	12	28
תיכונים בפריפריה	23	45	37	15

ברמת מובהקות 5%, האם טענת הצבא נכונה?

ב. במידגם בקורת של 200 חוזים נתגלו טעויות כדלקמן.

מספר טעויות לחוזה	0	1	2	3	4
מספר חוזים	75	70	40	10	5

קבע האם ברמת בטחון 95% התפלגות הטעויות היא פואסונית

2. לבדיקת מידת האהדה לשימוש באינטרנט ברחבי העולם גייס משרד פרסום עובדים על מנת לבצע סקר עולמי בקרב אוכלוסיות מדינות שונות. לצורך כך חולקה אוכלוסיית העולם לפי 5 יבשות. להלן הנתונים אודות כל יבשת:

יבשת	אחוז האוכלוסייה	עלות דגימה בודדת
אסיה	35%	25 ₪
אירופה	20%	4 ₪
אפריקה	10%	9 ₪
אמריקה	30%	64 ₪
אוסטרליה	5%	81 ₪

ידוע שהשונות של המשתנה הנדגם זהה באירופה, אמריקה ואוסטרליה. כמו כן ידוע ששונות המשתנה הנדגם הן באסיה והן באפריקה גבוהה פי 4 מהשונות באירופה. תקציב הסקר עמד על 2 מיליון ₪, ועלות הכנתו ללא תלות במספר הנדגמים עמדה על 150,000 ₪.

א. מצא מה מספר הנדגמים האופטימאלי בכל שכבה.

ב. משרד הפרסום התייעל, וכעת עלות הכנת הסקר ללא תלות במספר הנדגמים עמדה על 50,000 ₪ בלבד. מבלי לחשב, הסבר האם יחול שינוי במספר הנדגמים האופטימאלי בכל שכבה, וכן ביחס בין הנדגמים בכל יבשת.

ג. מבלי לחשב, כיצד תשתנה תשובתך לגבי היחס בין הנדגמים בכל יבשת, אם:

a. ידוע שהאמריקאים אחידים בדעתם, כלומר סטיית התקן של המשתנה הנדגם

באמריקה נמוכה מאד?

b. עלות הדגימה הבודדת באירופה זולה באופן מהותי מאד מעלות הדגימה בשאר

היבשות?

ד. מהו גודלו של המידגם המינימלי שיש לבצע כדי שבהסתברות 95% לא יסטה אחוז אוהדי השמוש באינטרנט במידגם ביותר מ-5% מאשר באוכלוסיה כולה?

3. א. נתון קובץ ציוני בחינה בסטטיסטיקה, אשר ממנו יצרו תרשים ענף עלה:

3	1
4	0
5	0 2
6	0115358
7	00320751
8	0334711
9	52401

(I) תן תמצית 5 נתונים וחווה דעתך על קובץ הנתונים מבחינת הסימטריה שלו. הסבר קביעתך.

(II) קבע האם יש חריגים, חריגים ביותר בנתוני הקובץ ונמק קביעתך.

(III) מה היתרונות בתרשים ענף עלה?

ב. רוכב אופניים נמצא במרחק מסויים מן היעד אליו עליו להגיע. הוא מתחיל לנוע לעבר היעד, כאשר הטבלה הנתונה מראה את מרחקו בק"מ מן היעד בזמנים שונים מרגע התחלת התנועה.

זמן בשעות	0.5	1	2	4	6	7	9	10
מרחק בק"מ מהיעד	115	105	90	70	60	50	30	10

מצא משוואת ישר באמצעות חציונים לאחר חלוקה לקבוצות. יש לקבוע נטייה ורמה ולבצע עליהם תיקון באמצעות השאריות

האם תוכל מתוך משוואת הקו להעריך את מהירות נסיעתו של הרכב ואת מרחקו מן היעד בטרם החל ברכיבתו ?

4. א. להלן נתוני רווח רבעוניים במאות אלפי ₪ של חברה:

1986		1985				1984			
4.2	4.8	3.1	4.1	3.3	3.5	3.0	3.1	2.8	3.1

תן שרטוט סכימטי לתהליך והתייחס לעונתיות ומגמה. מצא בהנחת עונתיות חיזוי ל-2 הרבעונים האחרונים בשנת 1986. פרט והסבר את חישוביך

ב. השתמש באותם נתונים וללא הנחת עונתיות לאפיין את התהליך כתהליך סטוכסטי (העזר בהפרשים) ומצא חיזוי לרבעון השלישי של 1986.

שיטות סטטיסטיות - ד"ר סניך למסמך א' תשע"א מועד א' פתרון -

פתרון שאלה 1

השתלבות הסטודנטים בפועל

סך הכל	שרון	יחידות ח"ר	יחידות שרון	יחידות ויס	חיל אויר ויס
90	11	29	15	35	תכנים בערים
70	7	40	9	14	המגזר הדרוזי
65	28	12	20	5	ישיבות תיכונות
120	15	37	45	23	תיכנים בפריפריה
345	61	118	89	77	סך הכל

חישוב הצפוי

סך הכל	שרון	יחידות ח"ר	יחידות שרון	יחידות ויס	חיל אויר ויס
90	15.91	30.78	23.22	20.09	תכנים בערים
70	12.38	23.94	18.06	15.62	המגזר הדרוזי
65	11.49	22.23	16.77	14.51	ישיבות תיכונות
120	21.22	41.04	30.96	26.78	תיכנים בפריפריה
345	61	118	89	77	סך הכל

חישוב - (צפו פחות נצפה) ברבע. חלקי הצפוי:

סך הכל	שרון	יחידות ח"ר	יחידות שרון	יחידות ויס	חיל אויר ויס
90	1.52	0.10	2.91	11.07	תכנים בערים
70	2.34	10.77	4.54	0.17	המגזר הדרוזי
65	23.71	4.71	0.62	6.23	ישיבות תיכונות
120	1.82	0.40	6.37	0.53	תיכנים בפריפריה

H0 אין תלות
H1 יש תלות

סכום כל הסטות: $\chi^2 = 77.82$
קיימות 9 דרגות חופש $\chi^2_{(9df)} = 16.92$

כלומר קיימת תלות.

0-30

סעיף ב:

		תוחלת הפגומים (יש לחשב ממוצע משוקלל)		סך הכל											
H0	H1	פואסון	אינה פואסון	1	200	4	5	3	10	2	40	1	70	0	75
						0.019		0.061		0.184		0.368		0.368	
					200	3.80		12.26		36.79		73.58		73.58	

הפרש: (בפועל - צפוי) גריבווע,

חלקי צפוי

סך ההפרש: 1.28

7.81

$\chi^2_{(3,0.05)} =$ קיימות 3 דרגות חופש
חמש מחלקות פחות 1 פחות 1 בשל
אומדן רשתות

לכן - H0, ההתפלגות פואסונית.

אפשר לאחד קבוצות כלל צפוי קטן מ- 5, אבל זה לא נחוץ

זמן שאלה 2

שנה	אחוז האנלסיה	עלות דגימה בודדת	שונות	סטיות תקן
בש	35%	25	4V	2σ
זיה	20%	4	V	σ
נירפה	10%	9	4V	σ
פרקה	30%	64	v	2σ
נוטרליה	5%	81	V	σ

זן התקצוב נטו:

$2,000,000 = 150,000 + 1,850,000$

$$\frac{n_{\text{אס}}}{n} = \frac{35\% \cdot 2\sigma / \sqrt{25}}{35\% \cdot 2\sigma / \sqrt{25} + 20\% \cdot \sigma / \sqrt{4} + 10\% \cdot 2\sigma / \sqrt{9} + 30\% \cdot \sigma / \sqrt{64} + 5\% \cdot \sigma / \sqrt{81}}$$

באופן דומה אירופה, אפריקה, אמריקה ואוסטרליה

חלק אסיה	חשוב	באחוזים
חלק אירופה	0.14	40.03%
חלק אפריקה	0.1	28.59%
חלק אמריקה	0.066667	19.06%
חלק אוסטרליה	0.0375	10.72%
	0.005556	1.59%
	0.349722	100.00%

המשוואה:

$1,850,000 = 25 \cdot X \cdot 40.03\% + 4 \cdot X \cdot 28.59\% + 9 \cdot X \cdot 19.06\% + 64 \cdot X \cdot 10.72\% + 81 \cdot X \cdot 1.59\%$

$1,850,000 = 21,0152 \cdot X$

$X = 88.032$

מספר האזרחים לבדיקה:

חלקת האזרחים לפי יבשות:

מחשב	מעגל
אסיה	27735.61
אירופה	19611.15
אפריקה	13207.43
אמריקה	7429.18
אוסטרליה	1100.62
סה"כ	69284.00

אָעף ב' זיפּר העדגמס בכל שכבה יעלה כי התקציב יגדל, היחס בין העדגמים בכל יבשת לא ישתנה כי עלות הדגימה בכל יבשת נשמרת זהה. באופן כללי מספר העדגמים יעלה נוך שמירה על היחס בין העדגמים בכל יבשת.

אָעף ג'

סעף 1: אם סטיית התקן אצל האמריקאים שואפת לאפס, החלק שלהם במדגם יהיה מינימלי וישאף אף הוא לאפס (למעשה אדם אחד ייצג את אמריקה). החלק של יותר היבשות יגדל כאשר היחס ביניהן יישמר למעשה שכן המונה יישאר כמו קודם והמכנה יקטן בגלל אמריקה.

סעף 2: אם עלות הדגימה באירופה זולה מאד, החלק היחסי שלהם במדגם יגדל מאד (המכנה קטן והמונה גדול). החלק של יותר היבשות יקטן כאשר היחס ביניהן יישמר למעשה שכן המונה יישמר כמו קודם והמכנה יגדל בגלל אירופה

וישוב מודל המידגם

הנחה שהאנליסיה גדולה מאוד מניח את גורם התיקון לאנליסיה סופית. מחסור ידע מוקדם על מידת האחדה לישראל ניחא 50% אחרים. כלומר התפלגות בינומית עם $P=0.5$ לפיכך $n=z^2pq/d^2$

$$n = 268.96$$

יש לדגום 269 איש

וערה: נלקחה רמת מובהקות 5% חד צדדית. גם פתרון דו צדדי יתקבל ($n=385$ איש)

פתרון שאלה 3

רשימת כמות המכונות ממוינת:

דרגה	כמות
1	31
2	40
3	50
4	52
5	60
6	61
7	61
8	63
9	65
10	65
11	68
12	70
13	70
14	70
15	71
16	72
17	73
18	75
19	77
20	80
21	81
22	81
23	83
24	83
25	84
26	87
27	90
28	91
29	92
30	94
31	95

חציון - האיבר במקום ה-(31+1) חלקי 2: כלומר האיבר ה-16
לכן החציון הוא: 72

רבעון ראשו - האיבר במקום ה-(31+1) חלקי 4, כלומר האיבר ה-8:
הרבעון הראשון: 63

הרבעון השלישי - במקום ה-24:
הרבעון השלישי: 83

א. תמצית חמישה נתונים:

דרגה	עומק	ציון	
31	1	95	↑
24	8	83	H2
16	16	72	M
8	8	63	H1
1	1	31	↓

ב. בדיקת סימטריה:

ממוצע רבעון 1 ורבעון 3:
ממוצע 73 חציון 72

ממוצע קצוות:
63 72

מאחר שקיים הבדל בין הממוצעים לחציון, הקובץ אינו סימטרי

ג. איתור חריגים:

$$S=1.5*(H1-H2)=1.5*20=30$$

ולכן הטווח של הנתונים שאינם חריגים: (H2+S, H1-S)
כלומר:

$$(63-30, 83+30) \rightarrow (33, 100)$$

ישנו חריג אחד: 31

איתור חריגים מאד:

מחוץ לטווח של (H2+2S, H1-2S), כלומר שאין חריגים ביותר

יתרונות הצגת ענף עלה: הצגה מסודרת של הנתונים ליצירת תמצית נתונים. ההצגה היא מעין היסטוגרמה ומאפשרת השוואה
back to back

תיקון רמה

E-(-0,625)	-0,625X	E=B0-114,375)	B0= -9,375-Y	-9,375X	Y	X
		5.3125	119.6875	-4.6875	115	0.5
0.625	-0.625	0	114.375	-9.375	105	1
		-5.625	108.75	-18.75	90	2
		-6.875	107.5	-37.5	70	4
0.625	-3.125	-2.5	111.875	-46.875	65	5
		1.875	116.25	-56.25	60	6
		1.25	115.625	-65.625	50	7
5.625	-5.625	0	114.375	-84.375	30	9
		-10.625	103.75	-93.75	10	10

-9.375	לוקחים חציין	$(105-30)/(1-9)$ $(105-65)/(1-5)$ $(65-30)/(5-9)$	-9.375 -10 -8.75	בדיקת שיפועים
--------	--------------	---	------------------------	---------------

$$y = -9,375x + 114,375$$

קו בסיסי

יש צורך בתיקון (סוף האיטרציות כאשר כל חציוני E מתאפסים). הקצוות הם אפסים לכן התיקון יתבצע
בין שני חציונים אחרים למשל

$(-2,5-0)/(5-1)$	-0.625	תיקון נטייה
	0.625	תיקון רמה

$$y = -10x + 115$$

קו מתוקן

מהירות נסיעתו של הרכב הוא שיפוע הקו כלומר 9,375 ק"מ בשעה. מרחקו מן היעד בטרם החל בנסיעה כלומר $X=0$ יהיה 115

פתרון סעיף א

ΔX_{t+1}	ביקת הפרשים (ΔX_t)	ערך מחכה DUI: מדד:	מדד ממומל (ראו חישוב למטה)	מדד עונתי: ממומל MA ממומל MA	ממומל MA	נתני רווח (DI)	רבעון	שנה
-0.008	-0.008	2.913	1.064	-	-	3.1	1	1984
-0.055	-0.055	2.905	0.964	-	-	2.8	2	1984
0.543	-0.055	2.850	1.088	1.016	3.050	3.1	3	1984
-0.104	0.543	3.393	0.884	0.949	3.163	3	4	1984
0.135	-0.104	3.289	1.064	1.045	3.350	3.5	1	1985
0.345	0.135	3.424	0.964	0.946	3.488	3.3	2	1985
-0.263	0.345	3.769	1.088	1.119	3.663	4.1	3	1985
1.004	-0.263	3.506	0.884	0.787	3.938	3.1	4	1985
-0.153	1.004	4.510	1.064	-	-	4.8	1	1986
0.431	-0.153	4.358	0.964	-	-	4.2	2	1986
0.017	0.431	4.789	1.088	-	-	5.210	3	1986
	0.017	4.806	0.884	-	-	4.249	4	1986

ראו חישוב הרגרסיה למטה

לאחר חישוב הרגרסיה

לאחר חישוב הרגרסיה

חישוב המדד:
1 רבעון 1.045
2 רבעון 0.946
3 רבעון 1.068
4 רבעון 0.868
סך הכל 3.927

מדד ממומל
1.064
0.964
1.088
0.884

הגרסאות הפרטים (על החלק המסומן למעלה):

XY	Y^2	X^2	Y	X
0.000	0.003	0.000	-0.055	-0.008
-0.030	0.295	0.003	0.543	-0.055
-0.057	0.011	0.295	-0.104	0.543
-0.014	0.018	0.011	0.135	-0.104
0.047	0.119	0.018	0.345	0.135
-0.091	0.069	0.119	-0.263	0.345
-0.264	1.008	0.069	1.004	-0.263
-0.153	0.023	1.008	-0.153	1.004
סכום	-0.562	1.547	1.524	1.597
ממוצע	-0.070	0.193	0.190	0.182
			0.182	0.200

$$S_{xy} = -0.562 - 8 \cdot 0.2 \cdot 0.182 = -0.8517793$$

$$S_{xx} = 1.524 - 8 \cdot 0.2^2 = 1.20487791$$

$$B1 = S_{xy}/S_{xx} = -0.71$$

$$B0 = Y_{avg} - B1 \cdot X_{avg} = 0.323$$

$$Y = -0.71X + 0.323$$

חישוב רבוען שליש ורבוען רביעי (מנכי מדד) - ראו בטבלה למעלה. בהדגשה

לאחר מכן מכפילים את התוצאה במדד ומקבלים את החיזוי לרבוען השלישי והרביעי בהתחשב במדד - ראו בטבלה למעלה בהדגשה

סעיף ב - מבלי להתייחס לעונתיות

שנה	רבעון	נתוני רווח	בדיקת ההפרשים (ΔX_t)	ΔX_{t+1}
1984	1	3.1	-	-0.3
1984	2	2.8	-0.3	0.3
1984	3	3.1	0.3	-0.1
1984	4	3	-0.1	0.5
1985	1	3.5	0.5	-0.2
1985	2	3.3	-0.2	0.8
1985	3	4.1	0.8	-1
1985	4	3.1	-1	1.7
1986	1	4.8	1.7	-0.6
1986	2	4.2	-0.6	0.917
1986	3	5.117	0.917	

ראו חישוב הרגרסיה להלן

התוצאה החזויה

רגרסיית הפרשים (על החלק המסומן למעלה):

XY	Y ²	X ²	Y	X	
-0.09	0.09	0.09	0.3	-0.3	
-0.03	0.01	0.09	-0.1	0.3	
-0.05	0.25	0.01	0.5	-0.1	
-0.10	0.04	0.25	-0.2	0.5	
-0.16	0.64	0.04	0.8	-0.2	
-0.80	1.00	0.64	-1.0	0.8	
-1.70	2.89	1.00	1.7	-1.0	
-1.02	0.36	2.89	-0.6	1.7	
סכום	-3.95	5.28	5.01	1.40	1.70
ממוצע	-0.49	0.66	0.63	0.18	0.21

$$S_{xy} = -3.95 - 8 \cdot 0.21 \cdot 0.18 = -4.2475$$

$$S_{xx} = 5.01 - 8 \cdot 0.21^2 = 4.64875$$

$$B_1 = S_{xy}/S_{xx} = -0.914$$

$$B_0 = Y_{avg} - B_1 \cdot X_{avg} = 0.369$$

$$Y = -0.914X + 0.369$$