הגדרות

*שכיח mo

- ניתן לחישוב לכל סוגי המשתנים (איכותי וכמותי)
- לא מושפע מערכים קיצוניים
- הוספה או הורדה קבועה לכל הסדרה = שכיח משתנה.

*חציון me

- לא מושפע מערכים קיצוניים
- החציון מושפע מספר הערכים ולא מערכים עצמם פרט לערך האמצעי שקובע את ערך החציון.

*ממוצע $\bar{X}$

- מושפע מכל הערכים כולל ערכים קיצוניים
- ניתן לחישוב במשתנה כמותי (מספרים)
- הוספת ערך = כל ערך שנוסיף מעל = ממוצע עולה
  מתחת = יורד
  שווה = לא משתנה
- הורדת ערך = כל ערך שנוריד מהסדרה שגבוה = ירד
  שקטן = יעלה
  ששווה = לא ישתנה

*מדדי מיקום:

50%	חציון	me	$\frac{iN}{2}$
25%	רבעון	Qi	$\frac{iN}{4}$
10%	עשירון	Di	$\frac{iN}{10}$
1%	אחוזון	Pi	$\frac{iN}{100}$

*הוספה קבועה לכל הערך הסדרה
- מדדי המיקום = עולים

*הורדת תוספת קבועה
- מדדי המיקום = יורדים

*מדדי פיזור

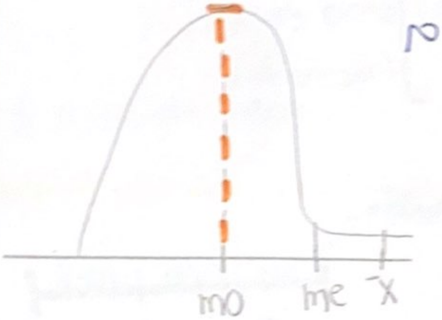
* R- כאשר מורידים או מוספים המרחק לא ישתנה
* S- 0 = כל הערכים שווים
* IOR- כל הוספת ערך קיצוני שגדול או קטן מהממוצע S=↑
  כל הורדת ערך קיצוני שגדול או קטן מהממוצע S=↓

איכותי (שמות)

- שמי
  - לא ניתן לדרג (אין משמעות לסדר)
  - * עוגה
- סדר
  - ניתן לדרג את הערכים
  - *מקלות

כמותי (מספרים)

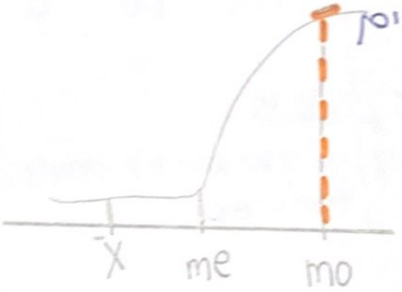
- בדיד
  - * מספרים בודדים
  - *מקלות
- רציף
  - *אין-סוף ערכים אפשריים
  - *היסטוגרמה.



*התפלגות א-סמטרית ימנית = ערכים נמוכים

$$\overset{\text{חציון}}{mo} < \overset{\text{שכיח}}{me} < \overset{\text{ממוצע}}{\bar{X}}$$

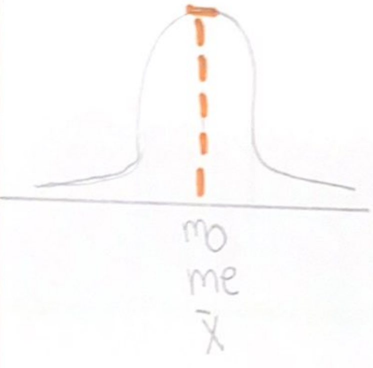
$$(mo=)\ me < \bar{X}$$



*התפלגות אי סמטרית שמאלית = ערכים גבוהים

$$\overset{\text{ממוצע}}{\bar{X}} < \overset{\text{שכיח}}{me} < \overset{\text{חציון}}{mo}$$

$$\bar{X} < me\ (=mo)$$



*התפלגות סימטרית

$$me = mo = \bar{X}$$

סיכום

מסגרות

* זמני ריצה - גבוה לוקח.

* מסגרת ציונים ↑

* ציוני תקן ופיזור יחסי:

ציון תקן. $Z_i = \leftarrow \frac{X_i - \bar{X}}{S}$

- מיקום יחסי
- חריג יותר
- טוב יותר → תלוי במסגרת

בנבדק יחיד. ← רחוק מ-0.

0 = ממוצע

פיזור יחסי $R_S = \leftarrow \frac{S}{X}$

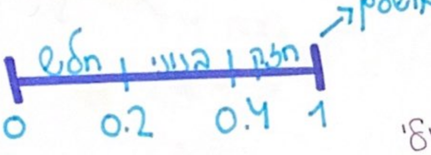
- פיזור יחסי
- פיזור
- הטרוגני = שונה (גבוה)
- הומוגני = דומה (נמוך)

#כללי ביחס לקבוצה. (מתייחס לכל הסדרה הסטטיסטית.)

* דומה יותר/ הומגני יותר/ הטרוגני פחות RS קטן

* שונה יותר/ הומגני פחות/ הטרוגני יותר RS גדול

* מדדי קשר - פירסון/ ספירמן/ קרמר:



קשר חיובי

$\uparrow X - \uparrow Y$

$\downarrow X - \downarrow Y$

דומה: עולים/ יורדים ביחד

קשר שלילי

$\uparrow X - \downarrow Y$

$\downarrow X - \uparrow Y$

שונים: אין קשר השני יורד והראשון עולה והפך.

מדד פירסון - Rp (כמותי · כמותי)

X_i	Y_i	$X_i - \bar{X}$	$y_i - \bar{y}$	$(X_i - \bar{X})^2$	$(y_i - \bar{y})^2$	$(x_i-\bar{x})(y_i-\bar{y})$
		0	0	סכום Σ	סכום Σ	סכום Σ

$$R_p = \frac{\sum^{\text{סכום}} (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2 \cdot \sum(y_i-\bar{y})^2}}$$

-1 = 1 ← תוצאה

בטווח

מדד ספירמן - Rs (סדר · סדר / סדר · כמותי)

X_i	y_i	R_x	R_y	$D = R_x - R_y$	D^2
					סכום Σ

$$R_S = 1 - \frac{6 \cdot \sum D^2}{N \cdot (N^2-1)}$$

מדד קרמר (שמ׳ · שמ׳ / שמ׳ · סדר)

$$E_i = \frac{\Sigma\text{שורה} \cdot \Sigma\text{טור}}{\Sigma\text{כללי}}$$ ערכים צפויים

$$X^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$

$$r_c = \sqrt{\frac{X^2}{N \cdot (L-1)}}$$ 5

L → * מספר הקן בין השורה לטור

שמות
ניתן
לדירוג

ספירמן (סדר X סדר / סדר X כמותי)

מספרים

* דניאל בחן את הקשר בין מידת ההשקעה של תלמיד ולבין הציון הקורס . להלן נתונים שהתקבלו

מידת ההשקעה X	ציון בקורס y
רבה	מצוין
מעטה	טוב
מעטה	חלש
בנוני	טוב
בנוני	חלש

האם קיים קשר בין המשתנים ?

שלב 1 הגדרת המשתנים

שלב 2 דירוג Rx Ry

שלב 3 בניית טבלה

שלב 4 חישוב עם הנוסחה

שלב 5 מסקנה

שלב 1

*הגדרת המשתנים: x בלתי תלוי ו y תלוי.

שלב 2

* דירוג המשתנים: Rx ו Ry.

Rx (לסדר לפי הסדר + מספור וחישוב.) Ry

מעטה 1, מעטה 2: $\frac{1+2}{2}=1.5$

בנוני 3, בנוני 4: $\frac{3+4}{2}=3.5$

רבה 5: 5

חלש 1, חלש 2: $\frac{1+2}{2}=1.5$

טוב 3, טוב 4: $3\frac{+4}{2}=3.5$

מצוין 5: 5

שלב 3: בניית טבלה:-

Xi	yi	Rx	Ry	D=Rx-Ry	D^2
רבה	מצוין	5	5	0	0
מעטה	טוב	1.5	3.5	-2	$(-2)^2=4$
מעטה	חלש	1.5	1.5	0	0
בנוני	טוב	3.5	3.5	0	0
בנוני	חלש	3.5	1.5	2	$2^2=4$

N=5

סה"כ 8

שלב 4: חישוב בנוסחה

$$r_s=1-\frac{6\cdot 8}{5\cdot(5^2-1}=0.6$$

$$\boxed{r_s=1-\frac{6\cdot\Sigma D^2}{N\cdot(N^2-1)}}$$

0 0.2 0.4 1

חלש בינוני חזק

שלב 5: מסקנה

תשובה: קיים קשר חיובי חזק

בדיד/ רציף

פירסון (כמותי X כמותי)

*מספרים.

- כדי לבדוק את היעילות במפעל מסוים נלקחו 6 עובדים שתיבדק הספק לוותק של העובד להספק יום עבודה

*שלב א׳: זיהוי y - x

משתנה תלוי. בלתי תלוי.

*שלב ב׳: חישובים ממוצעים.

$\bar{X} = \frac{36}{6} = 6$

$\bar{y} = \frac{110}{6} = 18.33$

כמותי הספק ביחידות מוצר y	כמותי וותק בשנים X
12	2
6	1
40	12
20	7
20	8
12	6~~8~~

שלב א - זיהוי משתנים

שלב ב - חישוב ממוצע

שלב ג - בניית טבלה

שלב ד - הצבה בנוסח

שלב ה - מסקנה

*שלב ג׳: בניית טבלה.

בלי מינוסים.

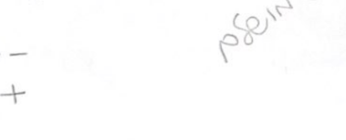
X_i	Y_i *נתון.	$X_i-\bar{X}$	$Y_i-\bar{Y}$	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$	$(X_i-\bar{X})(Y_i-\bar{Y})$
2	12	2-6=-4	-6.3333	16	40.0689	25.32
1	6	-5	-12.33	25	152.0289	61.65
12	40	6	21.67	36	469.5889	130.02
7	20	1	1.67	1	2.7889	1.67
8	20	2	1.67	4	2.7889	3.34
6	12	0	-6.33	0	40.0689	0
		0	0	82	707.28	222

(צריך להיות שווה=0)

*שלב ד׳: הצבה בנוסח-rp

$$rp = \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\Sigma(x_i-\bar{x})^2 \cdot \Sigma(y_i-\bar{y})^2}}$$

$$rp = \frac{222}{\sqrt{82 \cdot 707.28}} = 0.921$$



- שלילי

+ חיובי

*שלב ה׳: מסקנה.

קיים קשר חיובי חזק בין המשתנים

*שמות - לא ניתן לדרג. *שאות - ניתן לדרוג.

קרמר (שמי x שמי / שמי x סדר)

ין
יובי
שלילין

סבר -
לא ניתן
לדרג!

* בדוק האם קיים קשר בין השתייכות דתיים לבין סגנון בילוי

ריקודי עם	דיסקוטקים	סרטים	x | y →
30	60	110	דתיים
20	140	40	חילונים

שלב 1: הגדרת משתנים
X בלתי תלוי = שורות
y תלוי = טורים

שלב 2: בניית טבלה דו-ממדית
וחישוב סה"כ טור/שורה.

א . מה ניתן ללמוד לגבי סוג הקשר?

שלב 6: מסקנה
קיים קשר חזק.

סה"כ	ריקודי עם	דיסקוטקים	סרטים	x | y →
200	30 O_3 E_3 25	60 O_2 E_2 100	110 O_1 E_1 75	דתיים
200	20 O_6 E_6 25	140 O_5 E_5 100	40 O_4 E_4 75	חילונים
400	50	200	150	סה"כ

תמיד יהיה שווה הטור/שורה.

שלב 3: חישוב ערכים צפויים (E_i)

נוסחה

$$\boxed{E_i = \frac{\sum \text{טור} \cdot \sum \text{שורה}}{\sum \text{כללי}}}$$

$$E_1 = \frac{150 \cdot 200}{400} = 75$$

$$E_2 = \frac{200 \cdot 200}{400} = 100$$

$$E_3 = \frac{50 \cdot 200}{400} = 25$$

$$E_4 = \frac{150 \cdot 200}{400} = 75$$

$$E_5 = \frac{200 \cdot 200}{400} = 100$$

$$E_6 = \frac{50 \cdot 200}{400} = 25$$

שלב 4: חישוב בריבוע

נוסחה

$$\boxed{X^2 = \sum \frac{(O_i - E_i)^2}{E_i}}$$

$$X^2 = \frac{(110-75)^2}{75} + \frac{(60-100)^2}{100} \ldots = 66.66$$

שלב 5: חישוב R_c

$$r_c = \sqrt{\frac{X^2}{N \cdot (L-1)}}$$

*L: המספר הקטן בין מספר השורות למספר טורים.

$$r_c = \sqrt{\frac{66.66}{400 \cdot (2-1)}} = 0.408$$

$$L = \min(\underset{\text{שורות}}{2}, \underset{\text{טורים}}{3}) = 2$$

* רגרסיה לניארית

שלב 1 הגדרת משתנים

X משתנה בלתי תלוי (סיבה)

y משתנה תלוי (תוצאה) — מתייחסים לע-

*** הסבר:**

* חותך $a=$

• כאשר X אפס כלומר...

אזי $\frac{a}{y}$ הרווח הצפוי.

שלב 2 בניית טבלה

X	y	X^2	y^2	Xy

שלב 3 חישובים

* ממוצע

$$\bar{X}=\frac{\sum X_i}{N} \qquad \bar{y}=\frac{\sum y_i}{N}$$

* סטיית תקן

$$S_x=\sqrt{\frac{\sum X^2}{N}-\bar{X}^2} \qquad S_y=\sqrt{\frac{\sum y^2}{N}-\bar{y}^2}$$

* שונות משותפת

$$S_{xy}=\frac{\sum x_i y_i}{N}-\bar{X}\cdot\bar{y}$$

שונות משותפת

$b\cdot\frac{S_x}{S_y}$

קשר (r)

* במידה ואני רוצה לדעת האם קיים קשר בין המשתנים

שלב 4 חישוב האומדנים ובניית קו רגרסיה

חישוב b

↓

חישוב a

↓

בניית משוואה

שיפוע $b=\frac{S_{xy}}{S_{x^2}}$] הפרש

חותך $a=\bar{y}-b\cdot\bar{X}$

בניית משוואה

$$\hat{y}=a+b\cdot X$$

שלב 5 - חישוב מקדם המתאם של פירסון (בדק מידת הדיוק)

$$\underbrace{R_p=\frac{S_{xy}}{S_x\cdot S_y}}_{\text{דרך א'}}=\overbrace{b\cdot\frac{S_x}{S_y}}^{\text{דרך ב'}}$$

r חזק/ חלש וכו'.

⊛ חשב את המקדם והסבר את המשמעות

⊛

a חותך- כאשר x שווה לאפס.. אז הרווח הצפוי ___ $y=a$

b שיפוע- על כל עלייה של 1 בx
כלומר על כל עלייה של 1. אז הרווח צפוי
לעלות/
לרדת.

הסתברות

*ריבוע הקסם

סה"כ ↓	Ā	A	
סה"כ B	Ā∩B רק B	A∩B זוכים	B
סה"כ B̄	Ā∩B̄ אף אחד	A∩B̄ רק A	B̄
1	סה"כ Ā	סה"כ A	סה"כ →

$\boxed{\begin{array}{cc}2 & 1\\ 1 & 0\end{array}}$

(2 אופציות)

✻ מה הסיכוי... אם לפחות אחד ^{1,2} המאורעות התרחש?

חיתוך של כל המאורעות

$$P\left(A\cap B \,\middle|\, \begin{array}{l}A\cap B\\ \bar{A}\cap B\\ A\cap\bar{B}\end{array}\right) = \frac{A\cap B}{''}$$

✻ מה הסיכוי...

אם לכל היותר אחד ^{0,1} המאורעות התרחש?

$$P\left(\bar{A}\cap\bar{B} \,\middle|\, \begin{array}{l}\bar{A}\cap\bar{B}\\ \bar{A}\cap B\\ A\cap\bar{B}\end{array}\right) = \frac{A\cap B}{''}$$

✻ האם יש תלות?

$P(A) =$

$P(A|B) =$

*כן = יש תלות

*לא = אין תלות

שים לב שווה?

אם הם שווים אין תלות

אם הם שונים יש תלות

✻ מה הסיכוי אחד מהם זוכה?

(•או = חיבור.)

$$P(\underset{\text{סיכוי של A}}{A\cap\bar{B}}) + (\underset{\text{סיכוי של B}}{\bar{A}\cap B})$$

✻ מה הסיכוי שלפחות אחד מהם יזכה? (לפחות 1,2)

3 דרכים

1

$$P(\underset{\text{סיכוי של 2}}{A\cap B}) + P(\overset{\text{סיכוי של B}}{\bar{A}\cap B}) + P(\underset{\text{סיכוי של A}}{A\cap\bar{B}}) =$$

$$1 - P(\overset{\text{לא זוכים 2}}{\bar{A}\cap\bar{B}}) = \quad 2 \rightarrow \left(\text{דרך אשלים}\right)$$

3

$$P(\underset{\text{סיכוי של 2}}{A\cup B}) = P(\underset{\text{סכום A}}{A}) + P(\underset{\text{סכום B}}{B}) - P(\underset{\text{סיכוי של 2}}{A\cap B})$$

✻ מה הסיכוי שלכל היותר אחד מהם יזכה? (לכל היותר 0,1)

2 דרכים

1

$$P(\underset{\text{אף אחד}}{\bar{A}\cap\bar{B}}) + P(\overset{\text{רק B}}{\bar{A}\cap B}) + P(\underset{\text{רק A}}{A\cap\bar{B}}) =$$

$$1 - P(\overset{\text{סיכוי של 2}}{A\cap B}) \quad 2$$

✻ ידוע שקבוצה הראשונה ניצחה מה הסיכוי שהקבוצה השנייה גם תנצח?

$$P(\overset{\text{~~שאלה~~}}{B}|\overset{\text{ידוע}}{A}) = \frac{\overset{\text{חיתוך ידוע ושאלה}}{P(A\cap B)}}{\underset{}{\text{ידוע } A}}$$

אי - תלות בין מאורעות

תים = וגם -> כפל

או -> חיבור

$$P(A|B)=P(A)$$

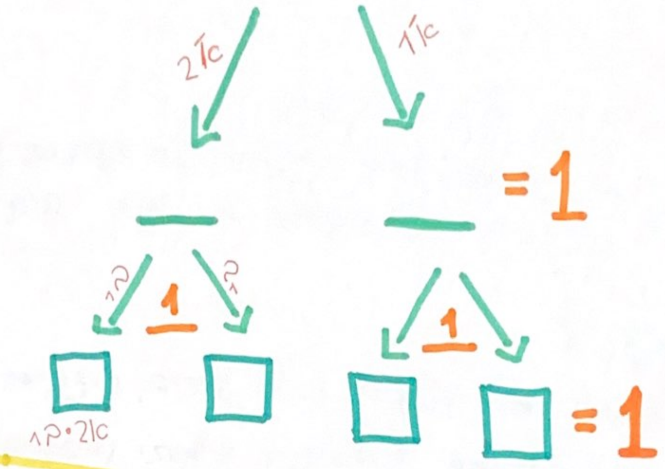
$$\downarrow$$

$$\frac{P(A\cap B)}{P(B)}=P(A)$$

$$P(A\cap B)=P(A)\cdot P(B)$$

אופן
מקרי?

$$\frac{1}{\text{מאורעות}}=\frac{1}{2}$$



* שתי קוביות הוגנות

*סכום גדול = לא כולל

*סכום קטן = לא כולל

מקסימום *סכום לכל היותר - כולל

מינימום *סכום לפחות - כולל