פתרון בחינה 1

חלק א' – שאלה 1

א) **לא נכון** החציון הוא הערך שמחצית מהמדגם קטן ממנו. הגדלת הערך הגדול ביותר לא משפיעה על החציון.

ב) **נכון** אם סטיית התקן שווה לאפס הרי שכל התצפיות שוות בערכן וממילא גם הרבעון התחתון שווה לרבעון העליון.

ג) **נכון** נחשב לפי ממוצע משוקלל ושונות מצורפת:

	גודל מדגם	ממוצע	שונות
מדגם מקורי	10	40	100
תוספת	2	$\frac{20+42}{2}=31$	$\frac{20^2+42^2}{2}-31^2=121$
סה"כ	12	$\bar{\bar{X}}=\frac{10\cdot40+2\cdot31}{12}=38.5$	$s_c^2=\frac{10\cdot100+2\cdot121}{12}+\frac{10\cdot(40-38.5)^2+2\cdot(31-38.5)^2}{12}=114.75$

ד) **נכון** ניסוי מקרי: הטלת קוביה פעמיים.

מרחב המדגם: $n(\Omega)=6^2=36$.

מאורע: $n(A)=9$

$$P(A)=\frac{9}{36}=\frac{1}{4}=0.25$$

ה) **לא נכון** ניסוי מקרי: לידת 4 ילדים (סיכוי שווה ללידת בן או בת).

מרחב המדגם: $n(\Omega)=2^4=16$

מאורע: A – במשפחה לפחות ילד אחד מכל מין (לפחות בן אחד ולפחות בת אחת).

המאורע המשלים: A^C – במשפחה כולם בנים או כולן בנות.

המאורע המשלים כולל שני מאורעות: B – כולם בנים ו-G – כולן בנות.

שני המאורעות B ו-G הם זרים ואיחודם הוא A^C. לכן נקבל:

$$P(A)=1-P(A^C)=1-P(B\cup G)=1-[P(B)+P(G)]=1-\left[\tfrac{1}{16}+\tfrac{1}{16}\right]=1-\tfrac{1}{8}=\tfrac{7}{8}$$

חלק ב'

שאלה 2

x מספר כוסות קפה בשבוע (אלפי ₪)	רוחב	אמצע	f(x)	F(x)	צפיפות
5-10	5	7.5	20	20	4
10-25	15	17.5	35	55	2.333
25-35	10	30	28	83	2.8
35-45	10	40	22	105	2.2
45-50	5	47.5	45	150	9

א) שכיח: אמצע המחלקה הצפופה ביותר $Mo=47.5$

חציון: לפי $F(x)$ מחלקת החציון היא המחלקה השלישית ולכן: $Md=25+\frac{75-55}{28}\cdot10=32.143$

ממוצע: $\bar{X}=\frac{7.5\cdot20+17.5\cdot35+...+47.5\cdot45}{150}=30.8$

ב) שונות וסטיית תקן: $s^2=\frac{7.5^2\cdot20+17.5^2\cdot35+...+47.5^2\cdot45}{150}-30.8^2=209.86\Rightarrow s=\sqrt{209.86}=14.486$

ג) $X\sim N(30.8,14.486^2)$

1. $P(X>40)=1-\Phi\left(\frac{40-30.8}{14.486}\right)=1-\Phi(0.63)=1-0.7357=0.2643$

26.43% שותים מעל 40 כוסות קפה בשבוע.

2. $z_{0.9}=1.282\Rightarrow X_{90\%}=30.8+1.282\cdot14.486=49.37$

שאלה 3

מחיר מכירה לפריט (ב-₪)	הסתברות
50	0.5
30	0.2
10	0.3

א) תוחלת: $E(X)=50\cdot 0.5+30\cdot 0.2+10\cdot 0.3=34$

שונות: $V(X)=50^2\cdot 0.5+30^2\cdot 0.2+10^2\cdot 0.3-34^2=304$

ב) הרווח הנקי ממכירת 500 פריטים: $Y=500X-10{,}000$:

תוחלת הרווח הנקי: $E(Y)=E(500X-10{,}000)=500E(X)-10{,}000=500\cdot 34-10{,}000=7{,}000$

שונות הרווח הנקי: $V(Y)=V(500X-10{,}000)=500^2V(X)=76{,}000{,}000$

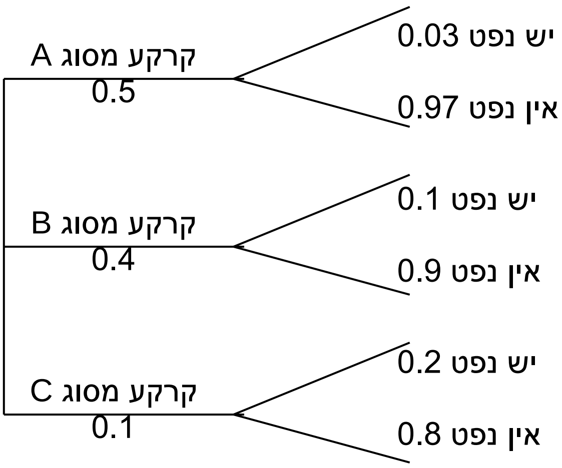
ג) המאורע: הרווח הנקי מכלל 500 הפריטים יהיה לפחות 5,000 ₪.

כלומר: הרווח הנקי מכל פריט יהיה לפחות 30 ₪ או $P(X\ge 30)$:

$$P(X\ge 30)=P(X=30)+P(X=50)=0.2+0.5=0.7$$

שאלה 4

נשרטט דיאגרמת עץ:



א) $P(\text{יש\_נפט})=0.5\cdot 0.03+0.4\cdot 0.1+0.1\cdot 0.2=$
$=0.015+0.04+0.02=0.075$

ב) $\dfrac{0.04}{0.075}=0.5333$

ג) 1. $P(C|B^C)=\dfrac{0.1}{1-0.4}=0.1667$

2. $P(\text{יש\_נפט}|B^C)=\dfrac{0.5\cdot 0.03+0.1\cdot 0.2}{0.6}=\dfrac{0.035}{0.6}=0.0583$

שאלה 5

א) נסמן: X – מנת משכל ו-Y – מספר צפיות במשך השבוע ונקבל:

$\sum x_i=970$ $\sum x_i^2=161{,}500$ $\sum y_i=19$ $\sum y_i^2=87$ $\sum x_iy_i=2{,}850$

נציב בנוסחת מקדם המתאם הלינארי: $r=\frac{6\cdot 2{,}850-970\cdot 19}{\sqrt{(6\cdot 161{,}500-970^2)(6\cdot 87-19^2)}}=\frac{-1{,}330}{2{,}126.99}=-0.6253$

ב) שני המשתנים בסולם סדר ולכן יש לחשב את מתאם ספירמן:

ערכים	השכלה	גבוהה	יסודית	גבוהה	גבוהה	תיכונית	תיכונית	יסודית
	מספר צפיות	6	6	1	4	3	4	2
דרגות	השכלה	6	1.5	6	6	3.5	3.5	1.5
	מספר צפיות	6.5	6.5	1	4.5	3	4.5	2
d^2	הפרש דרגות בריבוע	0.25	25	25	2.25	0.25	1	0.25

$$\Rightarrow D^2=0.25+1+0.25+2.25+25+25+0.25=54\Rightarrow r_s=1-\frac{6\cdot 54}{7\cdot(49-1)}=0.036$$

פתרון בחינה 2

חלק א' – שאלה 1

א) **לא נכון** $n(B)=13 \Rightarrow P(B)=\frac{13}{52}=\frac{1}{4}$ $n(A)=52-4=48 \Rightarrow P(A)=\frac{48}{52}=\frac{12}{13}$

המאורעות בלתי תלויים $n(A\cap B)=12 \Rightarrow P(A\cap B)=\frac{12}{52}=\frac{3}{13}=P(A)\cdot P(B)=\frac{12}{13}\cdot\frac{1}{4}=\frac{3}{13} \Rightarrow$

ב) **נכון** שני הצילומים הראשונים לא מוצלחים והשלישי מוצלח. ההסתברות לכך היא: $0.1^2\cdot 0.9=0.009$

ג) **נכון** זהו סולם סדר ולכן טרנספורמציה לינארית חיובית לא תשבש את סדר הערכים.

ד) **נכון** יש לבחור 2 בעלי תפקידים מבין 29 מועמדים. מספר האפשרויות לכך (לפי חוק המכפלה) הוא $29\cdot 28=812$

ה) **לא נכון** הממוצע המשוקלל חייב להיות בין שני הממוצעים של שתי הקבוצות. אם ערכו של הממוצע המשוקלל הוא 72 והממוצע של הסטודנטים לניהול וכלכלה הוא 78 הרי שהממוצע של הסטודנטים לחינוך ופסיכולוגיה צריך להיות נמוך מ-72.

חלק ב'

שאלה 2

א) אלו נתונים בדידים. נוסיף את עמודת השכיחות המצטברת:

מספר מקלטי TV x	מספר משפחות $f(x)$	$F(x)$
0	10	10
1	40	50
2	30	80
3	15	95
4	5	100

שכיח הערך בעל השכיחות הגבוהה ביותר ולכן $Mo=1$

חציון נתונים בדידים ולכן על-סמך גודל מדגם זוגי נקבל $Md=\frac{x_{(50)}+x_{(51)}}{2}=\frac{1+2}{2}=1.5$

ממוצע $\bar{X}=\frac{0\cdot 10+1\cdot 40+2\cdot 30+3\cdot 15+4\cdot 5}{100}=\frac{165}{100}=1.65$

אמצע הטווח $MR=\frac{x_{(1)}+x_{(n)}}{2}=\frac{0+4}{2}=2$

ב) שונות $s^2=\frac{0^2\cdot 10+1^2\cdot 40+2^2\cdot 30+3^2\cdot 15+4^2\cdot 5}{100}-1.65^2=1.028$

סטיית תקן $\Rightarrow s=\sqrt{1.028}=1.014$

ג) שכיח **ללא שינוי** גם אם כל 20 המשפחות בעלות 3 מכשירי טלוויזיה, הרי ששכיחות הערך 3 תהיה 35 ולא תעבור את שכיחות הערך 1 שהוא השכיח המקורי.

חציון **יגדל** גודל המדגם גדל ל-120 ולכן החציון הוא הממוצע של התצפית ה-60 וה-61 ושתיהן שוות ל-2.

ממוצע **יגדל** הוספת ערכים הגדולים מהממוצע מגדילה אותו.

אמצע הטווח **ללא שינוי** אין שינוי בערכים אלא רק בשכיחויות שלהם.

שונות וס. תקן **יגדלו** הוספת ערכים הרחוקים מהממוצע מגדילה את השונות ולכן גם את סטיית התקן.

שאלה 3

א) $0.7+0.3\cdot 0.4=0.7+0.12=0.82$

ב) $0.3\cdot 0.6\cdot 0.8=0.144$

ג) $\frac{0.3\cdot 0.4}{1-0.144}=\frac{0.12}{0.856}=0.14$

שאלה 4

$X\sim N\left(4.6,0.6^2\right)$ הישגי קפיצה למרחק (מטרים)

א) $z_{95\%}=1.645\Rightarrow X_{95\%}=4.6+0.6\cdot 1.645=5.587$
המרחק המינימלי כדי לקבל תעודת ספורטאי מצטיין הוא 5.59 מטר.

ב) $P\left(X>4.75\right)=1-\Phi\left(\frac{4.75-4.6}{0.6}\right)=1-\Phi\left(0.25\right)=1-0.5987=0.4013$
40.13% מהתלמידים קפצו למרחק גדול יותר מ-4.75 מטר.

ג) $Y\sim B\left(3,0.4013\right)$ - מספר הקפיצות למרחק של לפחות 4.75 מטר.
$$P\left(Y\geq 1\right)=1-P\left(X=0\right)=1-0.5987^3=0.7854$$

ד) $T\sim N\left(12.6,1.1^2\right)$ - הישגי ריצת 100 מטר (שניות).

ציון התקן של יובל בריצה הוא $z_T=\frac{11.8-12.6}{1.1}=-0.7273$

ציון התקן של יובל בקפיצה למרחק הוא $z_X=\frac{4.9-4.6}{0.6}=0.5$

ההמלצה היא שיובל יתחרה בריצה.
ההישג שלו בריצה טוב יותר מאשר בקפיצה למרחק בהשוואה לכלל המתחרים בבית הספר (לאור העובדה שככל שזמן הריצה הוא קטן יותר כך ההישג טוב יותר).

שאלה 5

א) נסמן: Y – לחץ דם ו-X – מינון ונקבל:

$$\sum x_i=115\qquad \sum x_i^2=1,625\qquad \sum y_i=1,241\qquad \sum y_i^2=156,893\qquad \sum x_i y_i=13,410$$

נציב בנוסחת מקדם המתאם הלינארי: $r=\frac{10\cdot 13,410-115\cdot 1,241}{\sqrt{\left(10\cdot 156,893-1,241^2\right)\left(10\cdot 1,625-115^2\right)}}=-0.9222$

ב) יש לחשב את מקדמי קו הניבויים $\tilde{Y}=a+bX$:

$$\left.\begin{aligned} b&=\frac{rs_Y}{s_x}=\frac{COV(X,Y)}{s_X^2}=\frac{\frac{13410}{10}-\frac{115}{10}\cdot\frac{1241}{10}}{\frac{1625}{10}-\left(\frac{115}{10}\right)^2}=\frac{-86.15}{30.25}=-2.848\\ a&=\bar{Y}-b\bar{X}=124.1-\left(-2.848\right)\cdot 11.5=156.852\end{aligned}\right\}\Rightarrow \tilde{Y}=156.852-2.848X$$

ג) שונות הניבויים לקו הניבויים $\tilde{Y}=a+bX$:

$$s_{\tilde{Y}}^2=r^2\cdot s_Y^2=\left(-0.9222\right)^2\cdot\left[\frac{156893}{10}-\left(\frac{1241}{10}\right)^2\right]=245.347$$

פתרון בחינה 3

חלק א' – שאלה 1

א) **לא נכון** מאורע A תנור ראשון פועל, מאורע B תנור שני פועל.

$$P(A^C \cup B^C) = P(A\cap B)^C = 0.1 \Rightarrow P(A\cap B) = 0.9$$

נשבץ את הנתונים בריבוע הקסם:

	A	A^C	
B	0.9	0.05	0.95
B^C	0.03	0.02	0.05
	0.93	0.07	1

האם מתקיים $P(A\cup B)\overset{?}{=}0.9$:

$$P(A\cup B) = P(A)+P(B)-P(A\cap B) = 0.95+0.93-0.9 = 0.98 \neq 0.9$$

ב) **לא נכון** יש לחשב שונות מצורפת. לפני כן יש לחשב ממוצע משוקלל.

$$\bar{\bar{X}} = \frac{35\cdot 75+40\cdot 70}{35+40} = \frac{2625+2800}{75} = \frac{5425}{75} = 72.333$$

$$s_C^2 = \frac{35\cdot 10^2+40\cdot 8^2}{75} + \frac{35\cdot(75-72.333)^2+40\cdot(70-72.333)^2}{75} =$$

$$= \frac{3500+2560}{75} + \frac{248.889+217.778}{75} = 87.023 \neq 80.8$$

ג) **נכון** ניסוי מקרי: הטלת שלוש קוביות

מרחב המדגם: $n(\Omega) = 6^3$

מאורע: שלוש תוצאות זהות $n(A) = 6\cdot 1\cdot 1 = 6$

הסתברות: $P(A) = \frac{n(A)}{n(\Omega)} = \frac{6}{6^3} = \frac{1}{6^2} = \frac{1}{36}$

ד) **לא נכון** $X \sim B(10,0.45)$ - מספר הביטולים. צריך לחשב את $1-P(X=10)$:

$$1-P(X=10) = 1-0.45^{10} \neq P(X=1) = 10\cdot 0.45\cdot 0.55^9$$

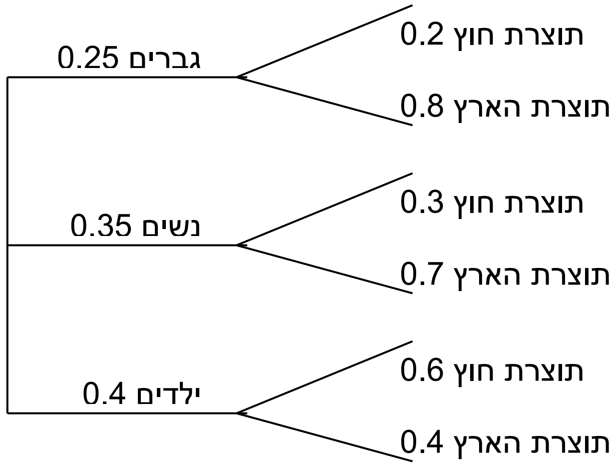
ה) **לא נכון** הפיזור יגדל היות ומוסיפים ערכים הרחוקים ממרכז ההתפלגות ובכך מגדילים את הפיזור. למשל, אם המשתנה הנמדד הוא בעל שלושה ערכים 1 , 2 ו-3 עם שכיחויות 10 , 20 ו-10 בהתאמה הרי שזו התפלגות סימטרית ומכאן שהממוצע הוא 2. השונות היא 0.5 לפי חישוב ממוצע הסטיות הריבועיות מהממוצע: $\frac{1^2\cdot 10+0^2\cdot 20+(-1)^2\cdot 10}{40} = \frac{20}{40} = 0.5$

לאחר הוספת שתי תצפיות, אחת שערכה 1 והשניה שערכה 3, הממוצע לא ישתנה וישאר 2. לכן גם ריבועי המרחקים מהממוצע לא ישתנו ויהיו בהתאמה 1 , 0 ו-1 אלא שעכשו משקלן של הסטיות הרחוקות מהממוצע יגדל ל-11 ולכן השונות תהיה: $\frac{11+11}{42} = \frac{22}{42} = \frac{11}{21} = 0.524$

חלק ב'
שאלה 2

7 נק'

נשרטט דיאגרמת עץ:



א) $P(\text{תוצרת\_הארץ}) = 0.25\cdot 0.8 + 0.35\cdot 0.7 + 0.4\cdot 0.4 =$
$= 0.2 + 0.245 + 0.16 = 0.605$

4 נק'

ב) $\dfrac{0.35\cdot 0.3}{1-0.605} = \dfrac{0.105}{0.395} = 0.2658$

8 נק'

ג) $X \sim B(10{,}0.605)$ - מספר מוצרים תוצרת הארץ.

צריך לחשב: $P(X\geq 2)$

$P(X\geq 2) = 1 - [P(X=0)+P(X=1)] =$

$1-\left[0.395^{10} + 10\cdot 0.605\cdot 0.395^{9}\right] = 0.9985$

6 נק'

ד) $Y \sim B(50{,}0.4)$ - מספר מוצרים עבור ילדים.

צריך לחשב: $E(Y), V(Y)$ $\qquad V(Y) = 50\cdot 0.4\cdot 0.6 = 12$ $\qquad E(Y) = 50\cdot 0.4 = 20$

שאלה 3

נוסיף את עמודת השכיחויות היחסיות:

x משכורת חודשית (אלפי ₪)	רוחב	אמצע	f(x)/n	F(x)/n	צפיפות יחסית
3-5	2	4	0.1	0.1	0.05
5-7	2	6	0.25	0.35	0.125
7-10	3	8.5	0.4	0.75	0.133
10-12	2	11	0.15	0.9	0.075
12-14	2	13	0.1	1	0.05

הערה: אפשר להניח גודל מדגם כלשהו (נניח 100) ולהמיר את עמודות השכיחות היחסית, השכיחות המצטברת היחסית והצפיפות היחסית לשכיחות, שכיחות מצטברת וצפיפות. אולם ניתן גם לענות על השאלה ללא המרה זו.

א) שכיח: אמצע המחלקה <u>הצפופה</u> ביותר $\qquad Mo = 8.5$

חציון: לפי $F(x)/n$ מחלקת החציון היא המחלקה השלישית ולכן: $\qquad Md = 7 + \frac{0.5-0.35}{0.4}\cdot 3 = 8.125$

ממוצע: $\qquad \bar{X} = 4\cdot 0.1 + 6\cdot 0.25 + 8.5\cdot 0.4 + 11\cdot 0.15 + 13\cdot 0.1 = 8.25$

ב) הטווח: $\qquad R = 14-3 = 11$

רבעון תחתון: לפי $F(x)/n$ מחלקת הרבעון התחתון היא מחלקה שניה $\qquad Q_1 = 5 + \frac{0.25-0.1}{0.25}\cdot 2 = 6.2$

רבעון עליון: קצה המחלקה השלישית, כי השכיחות המצטברת היא 0.75 $\qquad Q_3 = 10$

הטווח הבינרבעוני: $\qquad Q_3 - Q_1 = 10 - 6.2 = 3.8$

ג) $C_9 = \left[\dfrac{9-7}{3}\cdot 0.4 + 0.35\right]\cdot 100 = 61.67\%$ $\qquad C_6 = \left[\dfrac{6-5}{2}\cdot 0.25 + 0.1\right]\cdot 100 = 22.5\%$

$$\Rightarrow C_9 - C_6 = 61.67\% - 22.5\% = 39.167\%$$

שאלה 4

א) קו הניבויים עובר דרך מפגש הממוצעים ולכן מתקיים: $\bar{Y}=9.5=0.6\cdot\bar{X}+3.5$. מכאן מתקבל: $\bar{X}=10$

ב) $b=r\cdot\frac{s_y}{s_x}\Rightarrow r=b\cdot\frac{s_x}{s_y}=0.6\cdot\frac{4}{2.5}=0.96$

ג) $Y\sim N(9.5,2.5^2)$ - זמן הרכבה (דקות)

1. $P(Y\geq 12)=1-\Phi\left(\frac{12-9.5}{2.5}\right)=1-\Phi(1)=1-0.8413=0.1587$
15.87% מהמוצרים מרוכבים בזמן העולה על 12 דקות.

2. $z_{90\%}=1.282\Rightarrow x_{90\%}=9.5+1.282\cdot 2.5=12.705$
10% מהמוצרים מרוכבים ב-12.705 דקות או יותר.

שאלה 5

א) ישנם 8 (שמונה) מצבים אפשריים:

מצב	X מספר פותרים	גובה זכיה (אלפי ₪)	P - הסתברות
אף אחד לא פתר	0	0	$0.5\cdot0.4\cdot0.3=0.06$
רק יוסי ענה נכון	1	32	$0.5\cdot0.4\cdot0.3=0.06$
רק יעקב ענה נכון	1	64	$0.5\cdot0.6\cdot0.3=0.09$
רק ברוך ענה נכון	1	16	$0.5\cdot0.4\cdot0.7=0.14$
רק יוסי ענה לא נכון	2	80	$0.5\cdot0.6\cdot0.7=0.21$
רק יעקב ענה לא נכון	2	48	$0.5\cdot0.4\cdot0.7=0.14$
רק ברוך ענה לא נכון	2	96	$0.5\cdot0.6\cdot0.3=0.09$
כולם ענו נכון	3	112	$0.5\cdot0.6\cdot0.7=0.21$

לסיכום מקבל את פונקצית ההסתברות הבאה:

x	0	1	2	3
$P(x)$	0.06	0.29	0.44	0.21

ב) תוחלת סכום התשלום למתחרים:
$0\cdot0.06+32\cdot0.06+64\cdot0.09+16\cdot0.14+80\cdot0.21+48\cdot0.14+96\cdot0.09+112\cdot0.21=65.6$
הזכיינית תשלם למתחרים 65,600 ₪ בממוצע.

פתרון בחינה 4

חלק א' – שאלה 1

א) ***נכון***

$$P(A\cap B)=P(A\cup B)-P(A\cap B^C)-P(A^C\cap B)=0.58-0.18-0.28=0.12$$
$$P(A)=P(A\cap B^C)+P(A\cap B)=0.18+0.12=0.3$$
$$P(B)=P(A^C\cap B)+P(A\cap B)=0.28+0.12=0.4$$
$$P(A)\cdot P(B)=0.3\cdot 0.4=0.12=P(A\cap B)$$

ב) ***לא נכון*** אמנם הממוצע ללא שינוי כי ממוצע שתי התצפיות שנוספו שווה בדיוק לממוצע המקורי.
החציון למדגם המקורי הוא הממוצע בין התצפית ה-50 וה-51 כאשר יתכן שערכיהן בהתאמה הם 58 ו-82. כתוצאה מהוספת שני הערכים החציון החדש של המדגם יהיה הממוצע בין התצפית ה-51 והתצפית ה-52 שערכיהן בהתאמה 60 ו-82 ומכאן שהחציון ישתנה ויהיה 71.

ג) ***נכון*** אם הצילום השלישי הוא מוצלח, הרי ששני הראשונים היו לא מוצלחים. ההסתברות לכך היא $0.1^2\cdot 0.9=0.009$

ד) ***לא נכון*** מדידת שביעות הרצון נעשית בעזרת שני משתנים בסולם סדר ולכן המדד המתאים לבדיקת הקשר הוא ספירמן.

ה) ***נכון*** נחשב לפי ניסוי מקרי:
הניסוי: ימי הולדת לפי חודשים של ארבעה אנשים.

מרחב המדגם: $n(\Omega)=12^4$

המאורע: ימי הולדת בחודשים שונים: $n(A)=(12)_4=12\cdot 11\cdot 10\cdot 9$

הסתברות: $P(A)=\dfrac{n(A)}{n(\Omega)}=\dfrac{(12)_4}{12^4}=\dfrac{12\cdot 11\cdot 10\cdot 9}{12^4}=\dfrac{11\cdot 10\cdot 9}{12^3}$

חלק ב'

שאלה 2

א) קו הניבויים עובר דרך מפגש הממוצעים כלומר מתקיים: $\bar{Y}=a+b\bar{X}$.
לפיכך מתקיים: $\bar{Y}=\frac{4,000}{50}=80$ ולכן $80=10+1.4\bar{X}$ ומכאן מתקבל $\bar{X}=\frac{80-10}{1.4}=50$.

ב)
$$s_x^2=\tfrac{175,000}{50}-50^2=1,000\Rightarrow s_x=\sqrt{1,000}=31.62$$
$$s_y^2=\tfrac{420,000}{50}-80^2=2,000\Rightarrow s_y=\sqrt{2,000}=44.72$$
$$b=\frac{r\cdot s_y}{s_x}\Rightarrow r=\frac{b\cdot s_x}{s_y}=\frac{1.4\cdot 31.62}{44.72}=0.9899$$

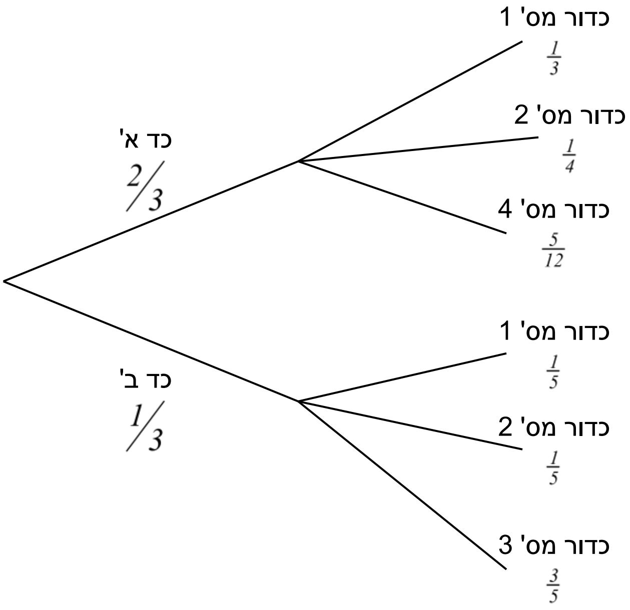
ג)
$$b'=\frac{r\cdot s_x}{s_y}=\frac{0.9899\cdot 31.62}{44.72}=0.7$$
$$a'=\bar{X}-b'\bar{Y}=50-0.7\cdot 80=-6$$

לכן קו הניבויים לניבוי X לפי Y הוא: $\tilde{X}=0.7Y-6\Rightarrow$

ד)
$$s_e^2=s_y^2-s_{\tilde{y}}^2=s_y^2-r^2\cdot s_y^2=s_y^2\cdot(1-r^2)=2,000(1-0.98)=40$$

<u>שאלה 3</u>

תיאור הניסוי בדיאגרמת עץ :



א) ההסתברות לכדור מס' 1 : $P(X=1)=\frac{2}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{1}{5}=\frac{2}{9}+\frac{1}{15}=\frac{13}{45}$

ההסתברות לכדור מס' 2 : $P(X=2)=\frac{2}{3}\cdot\frac{1}{4}+\frac{1}{3}\cdot\frac{1}{5}=\frac{1}{6}+\frac{1}{15}=\frac{7}{30}$

ההסתברות לכדור מס' 3 : $P(X=3)=\frac{1}{3}\cdot\frac{3}{5}=\frac{1}{5}$

ההסתברות לכדור מס' 4 : $P(X=4)=\frac{2}{3}\cdot\frac{5}{12}=\frac{5}{18}$

ב) $E(X)=1\cdot\frac{13}{45}+2\cdot\frac{7}{30}+3\cdot\frac{1}{5}+4\cdot\frac{5}{18}=\frac{37}{15}=2\frac{7}{15}=2.467$

$V(X)=1^2\cdot\frac{13}{45}+2^2\cdot\frac{7}{30}+3^2\cdot\frac{1}{5}+4^2\cdot\frac{5}{18}-\left(\frac{37}{15}\right)^2=\frac{311}{225}\Rightarrow\sigma_X=\sqrt{\frac{311}{225}}$

ג) $Y=10X-20\Rightarrow E(Y)=10E(X)-20=10\cdot\frac{37}{15}-20=4\frac{2}{3}$

$$V(Y)=10^2V(X)=100\cdot\frac{311}{225}=138\tfrac{2}{9}\Rightarrow\sigma_Y=11.76$$

<u>שאלה 4</u>

נתון : $X\sim N\left(2{,}000,s^2\right)$ משקל מחשבי מחברת (גרמים)

$$P(1{,}706\le X\le 2{,}294)=0.95$$

א) $0.95=P(1{,}706\le X\le 2{,}294)=P\left(-\frac{294}{s}\le Z\le\frac{294}{s}\right)\Rightarrow 0.975=P\left(X\le\frac{294}{s}\right)=\Phi\left(\frac{294}{s}\right)$

$$z_{97.5\%}=1.96=\frac{294}{s}\Rightarrow s=\frac{294}{1.96}=150$$

ב) $z_{10\%}=-z_{90\%}=-1.282\Rightarrow X_{10\%}=2{,}000-150\cdot 1.282=1{,}807.7$

10% מהמחשבים שוקלים פחות מ-1,807.7 גרם.

ג) $Y\sim B(5,p)$ - מספר מחשבי מחברת השוקלים יותר מ-2,150 גרם. צריך לחשב את $P(Y\le 2)$.

$$p=P(X\ge 2{,}150)=1-\Phi(1)=1-0.8413=0.1587\Rightarrow Y\sim B(5{,}0.1587)$$

$$P(Y\le 2)=\sum_{k=0}^{2}P(Y=k)=0.8413^5+5\cdot 0.1587\cdot 0.8413^4+\binom{5}{2}\cdot 0.1587^2\cdot 0.8413^3=$$

$$=0.4215+0.3975+0.15=0.969$$

שאלה 5

הנתונים בגבולות אמיתיים. להלן טבלה הכוללת את רוחבי המחלקות, אמצען ששכיחות מצטברת.

שעות צפיה גבולות אמיתיים	רוחב	אמצע	שכיחות $f(x)$	שכיחות מצטברת $F(x)$	צפיפות
0-10	10	5	120	120	12
10-16	6	13	150	270	25
16-20	4	18	100	370	25
20-24	4	22	120	490	30
24-30	6	27	70	560	11.67
30-40	10	35	40	600	4

א) ממוצע

$$\overline{X}=\frac{5\cdot120+13\cdot150+18\cdot100+22\cdot120+27\cdot70+35\cdot40}{600}=\frac{12,280}{600}=17.133$$

חציון מחלקה שלישית

$$Md=16+\frac{300-270}{100}\cdot4=17.2$$

שכיח אמצע המחלקה *הצפופה* ביותר

$$Mo=22$$

ב) רבעון תחתון מחלקה שניה

$$Q_1=10+\frac{150-120}{150}\cdot6=11.2$$

רבעון עליון מחלקה רביעית

$$Q_3=20+\frac{450-370}{120}\cdot4=22.667$$

טווח בינרבעוני

$$\Rightarrow Q_3-Q_1=22.667-11.2=11.467$$

ג) אחוז המשפחות שצופות עד 6 שעות בשבוע:

$$C_6=\left[\frac{6-0}{10-0}\cdot120+0\right]\cdot\frac{100}{600}=12\%$$

אחוז המשפחות שצופות עד 15 שעות בשבוע:

$$C_{15}=\left[\frac{15-10}{16-10}\cdot150+120\right]\cdot\frac{100}{600}=40.833\%$$

אחוז המשפחות שצופות בין 6 ל-15 שעות בשבוע:

$$\Rightarrow 40.833\%-12\%=28.833\%$$

מספר המשפחות שצופות בין 6 ל-15 שעות בשבוע:

$$\Rightarrow 28.833\%\cdot600=173$$

פתרון בחינה 5

חלק א' – שאלה 1

א) ***נכון*** לפי טבלת החיתוכים*:

	עישון		
	כן	לא	סה"כ
גברים	0.1	0.35	0.45
נשים	0.2	0.35	0.55
סה"כ	**0.3**	0.7	1

* הערכים עם הרקע האפור הם נתוני השאלה.
* הערך המובלט עם הרקע הצהוב הוא הנתון המבוקש בטענה.

ב) ***נכון*** אם X הוא משתנה מקרי הסופר תשלום של חברת הביטוח למבוטח במקרה מוות הרי שפונקצית ההסתברות שלו היא:

X תשלום למבוטח (אלפי $)	הסתברות
0	0.95
10	0.05

תוחלת התשלום למבוטח היא: $E(X)=0\cdot 0.95+10\cdot 0.05=0.5$
כלומר, חברת הביטוח תשלם לכל מבוטח בממוצע 500$.
כדי לשמור על איזון תקציבי על חברת הביטוח לגבות פרמיה בסכום זה.

ג) ***לא נכון*** ממוצע ההכנסה של כל הסטודנטים תלוי בשכיחות הסטודנטים לתואר ראשון ולתארים מתקדמים.

ד) ***לא נכון*** התוספת הזו תגדיל את השונות כי נוספים ערכים בקצוות של ההתפלגות ובכך הפיזור גדל.

ה) ***נכון*** נחשב לפי ממוצע משוקלל ושונות מצורפת:

	גודל מדגם	מחיר ממוצע	שונות המחירים
מלאי קיים	60	35	64
משלוח	20	35	0
סה"כ	80	35	$s_c^2=\frac{60\cdot 64+20\cdot 0}{80}+0=\frac{60\cdot 64}{80}=48$

אם הממוצעים זהים הממוצע המשוקלל שווה לממוצע של תתי המדגמים.
אם סדרת ערכים זהה בערכיה הרי שכל מדדי הפיזור שווים לאפס ולכן שונות המחירים של המשלוח היא אפס.
היות והממוצעים של המלאי של המשלוח זהים הרי ששונות הממוצעים (הרכיב השני של השונות המצורפת) הוא אפס.

חלק ב'

שאלה 2

$X\sim N(500,5^2)$ משקל כיכר לחם (גרם).

א) $z_{0.1}=-z_{0.9}=-1.282 \quad \Rightarrow \quad X_{10\%}=500-1.282\cdot 5=493.59$

ב) $P(492\le X\le 504)=\Phi\left(\frac{504-500}{5}\right)-\Phi\left(\frac{492-500}{5}\right)=\Phi(0.8)-\Phi(-1.6)=0.7881-1+0.9452=0.7333$

$\Leftarrow 0.7333\cdot 10{,}000=7{,}333$ ולכן במאפיה הזו נאפים 7,333 כיכרות לחם שמשקלם בין 492 גרם לבין 504 גרם.

ג) ההסתברות לכיכר לחם לא תקנית: $P(X\le 495)=\Phi\left(\frac{495-500}{5}\right)=\Phi(-1)=1-0.8413=0.1587$

(1) $Y\sim B(5,0.1587)$ מספר כיכרות לחם לא תקניים. מחפשים $P(Y\ge 1)$:

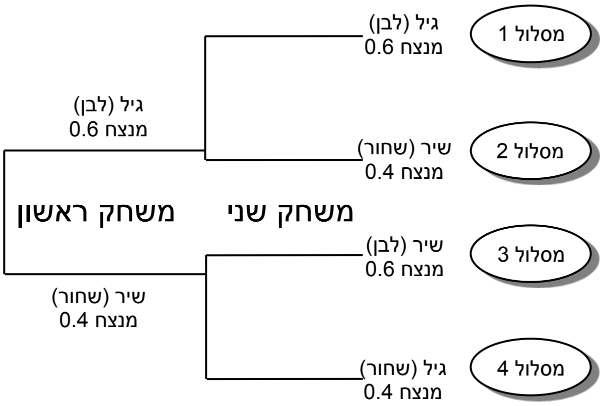
$P(Y\ge 1)=1-P(Y=0)=$
$=1-0.8413^5=0.5785$

(2) $T\sim B(200,0.8413)$ מספר כיכרות לחם תקניים:

$E(T)=200\cdot 0.8413=168.26$
$V(T)=200\cdot 0.8413\cdot 0.1587=26.7$

שאלה 3

תיאור הניסוי בדיאגרמת עץ:



א) מסלולים 2+3:
$$0.6\cdot 0.4+0.4\cdot 0.6=0.24+0.24=0.48$$

ב) מסלול 3 מתוך מסלולים 2+3:
$$\frac{0.4\cdot 0.6}{0.6\cdot 0.4+0.4\cdot 0.6}=\frac{0.24}{0.48}=0.5$$

ג) מסלול 1 מתוך מסלולים 1+3:
$$\frac{0.6\cdot 0.6}{0.6\cdot 0.6+0.4\cdot 0.6}=\frac{0.36}{0.36+0.24}=\frac{0.36}{0.6}=0.6$$

ד) טווח הערכים של המשתנה המקרי הסופר את מספר המשחקים ששיר ניצח הוא $\{0,1,2\}$ ופונקצית ההסתברות היא:

מספר משחקים ששיר ניצח	X	0	1	2
הסתברות	$P(X)$	0.36	0.4	0.24

$$\Rightarrow E(X)=0\cdot 0.36+1\cdot 0.4+2\cdot 0.24=0.88$$

שאלה 4

א) סוגי התרופות הוא משתנה בסולם שמי (נסמנו ב-X) ורמת ההקלה הוא משתנה בסולם סדר (נסמנו ב-Y). המשתנה הקובע הוא סוגי התרופות (הסולם הנמוך יותר) ולכן יש לחשב את מיתאם למדה ואת מיתאם קרמר.

נוסיף לטבלה את השכיחויות השוליות:

		תרופה X			
		א'	ב'	ג'	**סה"כ**
הקלה Y	אין	12	9	9	**30**
	מעטה	31	32	27	**90**
	רבה	7	9	14	**30**
סה"כ		**50**	**50**	**50**	**150**

מיתאם למדה:

$$L_X=150-50=100$$
$$L_{X|Y}=(30-12)+(90-32)+(30-14)=92$$
$$\Rightarrow \lambda_{X|Y}=1-\tfrac{92}{100}=0.08$$

$$L_Y=150-90=60$$
$$L_{Y|X}=(50-31)+(50-32)+(50-27)=60$$
$$\Rightarrow \lambda_{Y|X}=1-\tfrac{60}{60}=0$$

מיתאם קרמר:

טבלת שכיחויות צפויה (Expected):

		תרופה X			
		א'	ב'	ג'	**סה"כ**
הקלה Y	אין	10	10	10	**30**
	מעטה	30	30	30	**90**
	רבה	10	10	10	**30**
סה"כ		**50**	**50**	**50**	**150**

וטבלת ערכי χ^2:

	תרופה X		
	א'	ב'	ג'
אין	0.4	0.1	0.1
מעטה	0.033	0.133	0.3
רבה	0.9	0.1	1.6

$$\Rightarrow \chi^2=3.667\Rightarrow r_C=\sqrt{\tfrac{3.667}{150\cdot 2}}=0.111$$

ב) לחישוב מיתאם פירסון לקשר לינארי יש לחשב את הסכומים, סכומי הריבועים וסכום המכפלות:

								סה"כ
X	מינון במ"ג	0.8	0.5	0.7	1	0.4	0.2	**3.6**
Y	שעות שינה	10	6	9	8	5	4	**42**
X^2		0.64	0.25	0.49	1	0.16	0.04	**2.58**
Y^2		100	36	81	64	25	16	**322**
$X \cdot Y$		8	3	6.3	8	2	0.8	**28.1**

$$r = \frac{n\sum xy-(\sum x)(\sum y)}{\sqrt{[n\sum x^2-(\sum x)^2][n\sum y^2-(\sum y)^2]}} = \frac{6\cdot 28.1-3.6\cdot 42}{\sqrt{(6\cdot 2.58-3.6^2)\cdot(6\cdot 322-42^2)}} = \frac{17.4}{\sqrt{2.52\cdot 168}} = \frac{17.4}{20.576} = 0.8456$$

<u>שאלה 5</u>

א) נעשה שימוש בממוצע על-מנת לאתר את אמצע המחלקה האחרונה:

$$\overline{X} = 36.2 = \frac{10\cdot 9+25\cdot 9+35\cdot 14+50\cdot 11+x_5\cdot 7}{50} \Rightarrow 36.2\cdot 50 = 90+225+490+550+x_5\cdot 7 \Rightarrow x_5 = 65$$

אמצע המחלקה האחרונה הוא 65. בגלל שהגבול התחתון שלה הוא 60 הרי שהגבול העליון שלה הוא 70.
נשלים את טבלת השכיחות:

מספר ניתוקים לשעה	אמצע	רוחב	שכיחות	שכיחות מצטברת	צפיפות
0-20	10	20	9	9	0.45
20-30	25	10	9	18	0.9
30-40	35	10	14	32	1.4
40-60	50	20	11	43	0.55
60-70	65	10	7	50	0.7

ב) <u>חציון</u> (מחלקה שלישית 30-40) $Md = 30 + \frac{25-18}{14}\cdot 10 = 35$

<u>שכיח</u> (אמצע המחלקה הצפופה ביותר) $Mo = 35$

ג) <u>רבעון תחתון</u> (מחלקה שניה 20-30) $Q_1 = 20 + \frac{12.5-9}{9}\cdot 10 = 23.9$

<u>רבעון עליון</u> (מחלקה רביעית 40-60) $Q_3 = 40 + \frac{37.5-32}{11}\cdot 20 = 50$

<u>טווח בינרבעוני</u> $\Rightarrow Q_3 - Q_1 = 50 - 23.9 = 26.1$

ד) יש לבצע טרנספורמציה לינארית של $Y=2X$ כאשר X הוא מספר הניתוקים בשעות הלילה (המשתנה המקורי של המדגם) ו-Y הוא מספר הניתוקים בשעות היום. כל המדדים יוכפלו פי 2.