

עבודת גמר חלק 1 - הסתברות

על ידי : מיכאל טרושקין

1.

א. בערך $1/125$ מהלידות זה תאומים לא זהים ו- $1/300$ מהלידות זה תאומים זהים. לאלביס היה אח תאום שמת בלידה. מה ההסתברות שאלביס היה תאום זהה? (ניתן להניח שההסתברות להולדת בן ובת שווה ל- $1/2$).

נתון שלאליביס היה תאום שמת בלידה, ההסתברות שיהיה תאום הוא $1/125$, בנוסף $1/300$ לידות הם תאומים זהים.

נשים לב כי ההסתברות לתאום בלידה בהנחתן שיש תהום זהה הוא 100% כלומר נסמן:

H – having a tween
 E – Having a similar tween

$$P(H | E) = 1$$

$$P(H | !E) = 0$$

$$P(E) = 1/125$$

$$P(H) = 1/300$$

נרצה לחשב מהי ההסתברות של תאום זהה בהנחתן תאומים.

$$P(E | H) = P(H | E) * P(H) / P(E) = \frac{\frac{1}{300}}{\frac{1}{125}} = \frac{125}{300} = \frac{5}{12}$$

ב. יש שתי קערות של עוגיות. בקערה 1 יש 10 עוגיות שקדים ו-30 עוגיות שוקולד. בקערה 2 יש 20 עוגיות שקדים ו-20 עוגיות שוקולד. אריק בחר קערה באקראי ובחר ממנה עוגיה באקראי. העוגיה שנבחרה היא שוקולד. מה ההסתברות שאריק בחר את קערה 1?

נשים לב כי נתון לנו שנבחרה עוגיית שוקולד. היות ובקערה 1 יש 30 עוגיות שוקולד, ובקערה 2, יש 20 עוגיות שוקולד, נסיק שההסתברות לבחור בקערה 1, היא בדיוק $30/50$ כלומר 0.6.

בנוסחא אחרת אפשר להגיד שההסתברות לבחור עוגיית שוקולד דווקא מקערה 1, היא כמו ההסתברות שבחרתי בקערה 1, ואז בחרתי בעוגיית שוקולד, לחלק להסתברות שאני בעולם של עוגיות השוקולד.

כלומר:

E – chocolate cookie
 H_1 = plate 1

$$P(H_1 | E) = \frac{P(E | H_1) \cdot P(H_1)}{P(E)} = \frac{\frac{30}{40} \cdot \frac{1}{2}}{\frac{50}{80}} = \frac{30}{50} = \frac{3}{5}$$

ברור שזאת אותה תשובה אבל דרך טיפה שונה.

בשנת 1995 חברת M&M הוסיפה את הצבע כחול. לפני השנה הזו, התפלגות הצבעים

בשקית M&M נראית כך:

30% Brown, 20% Yellow, 20% Red, 10% Green, 10% Orange, 10%

Tan

החל משנת 1995, ההתפלגות נראית כך:

24% Blue, 20% Green, 16% Orange, 14% Yellow, 13% Red, 13% Brown.

לחבר שלכם יש 2 שקיות M&M, אחת משנת 1994 ואחת משנת 1996 והוא לא מוכן לגלות לכם איזו שקית שייכת לאיזו שנה. אבל הוא נותן לכם סוכריה אחת מכל שקית. סוכריה אחת היא צהובה ואחת היא ירוקה. מה הסיכוי שהסוכריה הצהובה הגיעה מהשקית של 1994?

תחילה נחשב את ההסתברות של להוציא סוכריה צהובה דווקא מהשקית של 1994 ושל 1996, נשים לב שההסתברות להוציא סוכריה משקית 1994 היא 20% ומשקית 2 היא 14% ולכן:

$$a(0.2 + 0.14) = 1$$

$$P(Y94) = \frac{0.2}{0.34} = 0.588$$

כלומר בהינתן שהוצאנו סוכריה צהובה, ההסתברות להציא סוכריה צהובה היא 100%, והיא בנויה מההסתברות להוציא סוכריה צהובה משקית 1 או שקית 2.

אבל זה לא הכל, כי יש לנו גם סוכריה ירוקה, וגם עבור הסוכריה הירוקה יש הסתברות כלשהי, אם הסוכריה הצהובה היא משקית 94 אז הירוקה היא מ 96.

ולכן ההסתברות שהוצאנו את הצהובה מ - 94, היא ההסתברות שהוצאנו צהובה מ 94 וירוקה מ 96. לחלק לכל האפשרויות להוציא סוכריה צהובה וירוקה משקיות שונות, כלומר, לחלק להסתברות שהירוקה מ 96 וצהובה מ 94, וגם ההסתברות שהירוקה מ 94 והצהובה מ 96.

ולכן: $H - Y94G96, E - GY$

$$P(H | E) = \frac{P(Y94) \cdot P(G96)}{P(Y94) \cdot P(G96) + P(Y96) \cdot P(G94)}$$

נחשב את שאר המשתנים:

$$P(Y94) = \frac{20}{34}, P(Y96) = \frac{14}{34}$$

$$P(G96) = \frac{2}{3}, P(G94) = \frac{1}{3}$$

ולכן נקבל:

$$P(H | E) = \frac{\frac{20}{34} \cdot \frac{2}{3}}{\frac{20}{34} \cdot \frac{2}{3} + \frac{14}{34} \cdot \frac{1}{3}} = \frac{\frac{20}{51}}{\frac{27}{51}} = \frac{20}{27} = 0.74074$$

כלומר ההסתברות שהסוכריה הצהובה מסכית 1, וסוכריה ירוקה מסכית 2 היא בערך 0.74074.

3. הלכת לדוקטור בעקבות ציפורן חודרנית. הדוקטור בחר בכך **באקראי** לבצע בדיקת דם הבדקת שפעת חזירים. ידוע סטטיסטית ששפעת זו פוגעת ב-1 מתוך 10,000 אנשים באוכלוסייה. הבדיקה מדויקת ב-99 אחוז במובן שההסתברות ל false positive היא 1%. הווה אומר שהבדיקה סיווגה בטעות אדם בריא כאדם חולה היא 1 אחוז. ההסתברות ל- false negative היא 0 – אין סיכוי שהבדיקה תגיד על אדם החולה בשפעת חזירים שהוא בריא. בבדיקה יצאת חיובי (יש לך שפעת).
א. מה ההסתברות שיש לך שפעת חזירים?

ב. נניח שחזרת מתאילנד לאחרונה ואתה יודע ש-1 מתוך 200 אנשים שחזרו לאחרונה מתאילנד, חזרו עם שפעת חזירים. בהינתן אותה סיטואציה כמו בשאלה א, מה ההסתברות (המתקנת) שיש לך שפעת חזירים?

אז נסמן $H - I am really sick$ ו $E - Positive test$. נחשב את $P(H | E)$.

$$P(H | E) = \frac{P(E | H) \cdot P(H)}{P(E)} = \frac{P(E | H) \cdot P(H)}{P(E | H) \cdot P(H) + P(E | !H) \cdot P(!H)}$$

לפי הנתון:

$$\begin{aligned} P(E | H) &= 1 \\ P(E | !H) &= 0.01 \\ P(H) &= \frac{1}{10000}, P(!H) = \frac{9999}{10000} \end{aligned}$$

אם כך נחשב את מה שהתבקשנו:

$$P(H | E) = \frac{\frac{1}{10000}}{\frac{1}{10000} + \frac{1}{100} \cdot \frac{9999}{10000}} = \frac{1}{1 + \frac{9999}{100}} = \frac{100}{10099} \approx 0.0099 \approx 0.99\%$$

כלומר קיבלנו שההסתברות להיות באמת חולה בהינתן שנבחרתי רנדומלית, ורק 1 מתוך 10000 חולה היא כמעט 1%.

ב. נעשה את אותו החישוב רק כשעט נשנה את הנתונים:

$$P(H) = \frac{1}{200}, P(!H) = \frac{199}{200}$$

ולכן:

$$P(H | E) = \frac{\frac{1}{200}}{\frac{1}{200} + \frac{1}{100} \cdot \frac{199}{200}} = \frac{100}{100 + 199} = \frac{100}{299} \approx 33\%$$

כלומר אם השכיחות של המחלה בהוכלוסייה שלי היא 1 מתוך 200 ויצאתי חיובי בטסט ההסתברות שאני חולה, היא בערך 33%.

Random Variables

1. Roi is playing a dice game with Yael.

Roi will roll 2 six-sided dice, and if the sum of the dice is divisible by 3, he will win 6\$. If the sum is not divisible by 3, he will lose 3\$.

What is Roi's expected value of playing this game?

the expected value of getting a sum divisible by 3, is getting the sum: 3, 6, 9, 12

hence the expected value is: Denote D – sum divisible, $!D$ – sum not divisible
would be $6 \cdot D3 + !D3 \cdot (-3)$

of course $!D3 = 1 - D3$, so lets calculate $D3$.

$D3$ is the sum of the probabilities getting 3, 6, 9, 12.

$$P(3) = \{1:2, 2:1\} = \frac{2}{36}, P(6) = \{1:5, 2:4, 3:3, 4:2, 5:1\} = \frac{5}{36}$$

$$P(9) = \{3:6, 4:5, 5:4, 6:3\} = \frac{4}{36}, P(12) = \{6:6\} = \frac{1}{36}$$

$$P(D3) = \frac{12}{36}$$

$$\text{hence the expected value is : } 6 \cdot \frac{12}{36} - 3 \cdot \frac{24}{36} = 0$$

2. Sharon has challenged Alex to a round of Marker Mixup. Marker Mixup is a game where there is a bag of 5 red markers numbered 1 through 5, and another bag with 5 green markers numbered 6 through 10.

Alex will grab 1 marker from each bag, and if the 2 markers add up to more than 12, he will win 5\$, 5. If the sum is exactly 12, he will break even, and If the sum is less than 12, he will lose 6\$.

the expected value of alex playing marker mixup:

is the probability of getting more then 12 times 5, plus the probability of getting less then 12 times 6.

note that the probability of getting less then 12 is 1 minus the probability of getting 12 and above.

the probability of getting 12 and above.

lets calculate it by calculating by saying he got X from bag A and what is the probability getting Y+.

lets calculate it by summing the number of way's to get 12+.

Bag 1: $1 \rightarrow 0$, $2 \rightarrow 0$, $3 \rightarrow 1$, $4 \rightarrow 2$, $5 \rightarrow 3$.

The total number of possibilities is : 25.

hence the probability of getting 12 + is : $\frac{6}{25}$.

the probability of getting exactly 12. is when we get : 2 : 10, 3 : 9, 4 : 8, 5 : 7

hence the probability is: $\frac{4}{25}$

then the expected value is:

$$\frac{6}{25} \cdot 5 + \frac{15}{25} \cdot (-6) = \frac{6}{5} - \frac{18}{5} = -2.4$$

3. A division of a company has 200 employees, 40%, percent of which are male. Each month, the company randomly selects 8 of these employees to have lunch with the CEO.

What are the mean and standard deviation of the number of males selected each month?

there is 80 males in the company.

each month, the company selects 8 random employees, selection 8 random, from a pool of 200 people when 40% are males would mean that 40% of those 8 are males.

hence on average there would be 3.2 selected each month, of course we can't select 3.2 males.

so we would assume that each 5 months 4 males selected and every other month 3 selected, on average.

assuming the 8 candidates are random we can assume, that on average every 4 months 40% of the chosen are males, hence from 40 people chosen every 5 months 16 would be males, with 3-4 males being chosen every month we would get the following set:

<i>Month</i>	1	2	3	4	5	<i>Total</i>
<i>Males</i>	3	3	3	3	4	16

with leads us to 3.2 mean.

now let's calculate the standard deviation:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2}$$

$\mu = 3.2, N = 5$

$$\sigma = \sqrt{\frac{0.16 + 0.64}{4}} \approx 0.4472$$

again so the mean males participating is 3.2 but we cannot take 3.2 males, so we would have to pick between 3 to 4 males each month, a reasonable sample of what would happen is the sample from above.

we can also assume this is actually what's happening and we will get STD:

$$\sigma = \sqrt{\frac{0.16 + 0.64}{5}} = 0.4$$

Different dealers may sell the same car for different prices. The sale prices for a particular car are normally distributed with a mean and standard deviation of 26,000\$ and 2,000\$, respectively. Suppose we select one of these cars at random. Let X = the sale price (in thousands of dollars) for the selected car.

Find $P(26 < X < 30)$,

$$\mu = 26, \sigma = 2$$

$$Z = \frac{x - \mu}{\sigma} = \frac{30 - 26}{2} = 2$$

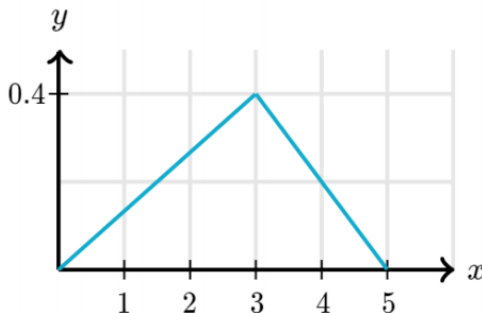
נמיר למצב שבוא הממוצע הוא 0, ונשתמש בנוסחא הבא:

$$\phi(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

הפונקציה הבאה היא ההסתברות שמישהו במרחק x (לפי סיגמה) מהממוצע. אצלנו אנחנו מסתכלים על ההסתברות להיות בתווח 0 עד 2, בפונקציה. ולכן על מה שעלינו לעשות זה לחשב את:

$$P(26 < x < 30) = \int_0^2 \phi(x) dx = 0.4772$$

5. Given the following distribution, what is $P(x > 3)$?



נחשב את השטח המשולש שמשמאל ל 3. נשים לב כי גובה המשולש הוא 0.4 והרוחב שלו הוא 2. ולכן:

$$P(x > 3) = \frac{0.4 \cdot 2}{2} = 0.4$$

היות וחישבנו את ההסתברות של $P(x \geq 3) = 0.4$ נסיק שההסתברות לממש גדול מ-3, יהיה כמעט שואף ל 0.4.

6. A company has 500 employees, and 60% of them have children. Suppose that we randomly select 4 of these employees.

What is the probability that exactly 3 of the 4 employees selected have children?

the company has 500 employees 60% of them have children hence $\frac{300}{500}$ have children,
the probability of having 3 out of 4 randomly chosen have children is then:

$$\frac{300}{500} \cdot \frac{299}{499} \cdot \frac{298}{498} \cdot \frac{200}{497} = 0.086$$

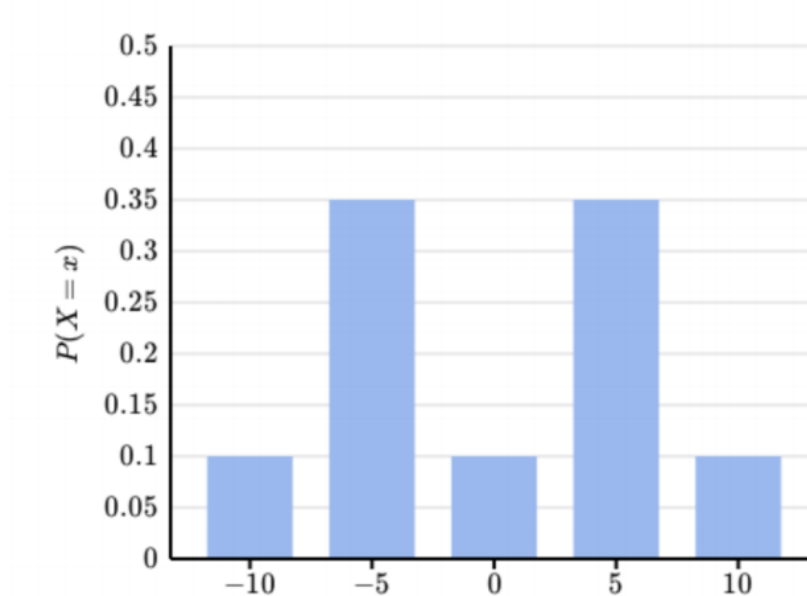
this is the probability of 3 out of 4 being with children but there is 4 ways to get this result.
either the first one without children or the second one, of the third one, or the forth one.

so in reality we will get this figure:

$$\frac{300}{500} \cdot \frac{299}{499} \cdot \frac{298}{498} \cdot \frac{200}{497} \cdot \binom{4}{1} = 0.34629$$

from a simple checkout we can see that by this formula we get a probability of 1,
if we combine 0 out of 4 1 out of 4 etc...
so we can see that this is indeed the correct formula.

7. Look at the next Graph. What is the expected value of X?



well the graph is simetrical and the mean is 0, so we can conclude that the expected value is the mean, hence, the expected value of x is 0.