הסתברות 1 אוניברסיטת אריאל

נכתב ע"י אייל לוי github.com/LeviEyal/Summaries

נוסחאות בסיסיות וכלליות:

- $P(\Phi)=0, P(\Omega)=1$
- $P(A^c)=1-P(A)$. (מאורע משלים)
- אם $A\subseteq B$ אז $P(A)\le P(B)$.
- אם A,B זרות אז: $P(A\cup B)=P(A)+P(B)$.
- $P(A\cap B)=P(A)\cdot P(B|A)=P(B)\cdot P(A|B)$
- $P(A\cup B|C)=P(A|C)+P(B|C)-P(A\cap B|C)$
- $P(A|B)=1-P(\bar{A}|B)$
- $P(A\cap B\cap C)=P(A)\cdot P(B|A)\cdot P(C|A\cap B)$
- $C\subseteq B\subseteq A \Rightarrow P(C)=P(A)\cdot P(B|A)\cdot P(C|B)$

קומבינטוריקה:

בחירת K איברים מ-N איברים:

	ללא החזרה	עם החזרה
ללא חשיבות לסדר	$\binom{n}{k}=\frac{n!}{k!\,(n-k)!}$	$\binom{n+k-1}{n-1}$
עם חשיבות לסדר	$\frac{n!}{(n-k)!}$	n^k

סידור N איברים **במעגל**: $(n-1)!$, **בשורה**: $n!$

הכלה והדחה:

$P(\bigcup_{i=1}^n A_i)=\sum_{i=1}^n P(A_i)-\sum_{i<j=1}^n P(A_i\cap A_j)+\sum_{j<k=1}^n P(A_i\cap A_j\cap A_k)-\cdots(-1)^{n+1}P(\bigcap_{i=1}^n A_i)$

התפלגות	משמעות	פונקציית התפלגות	תוחלת $E(X)$	שונות $Var(X)$
ברנולי $\mathbf{Ber(p)}$	ראה אינדיקטורים	$\begin{cases}p, & x=1\\ q, & x=0\end{cases}$	p	$p\cdot q$
אחידה $\boldsymbol{U(a,b)}$	לכל איבר בטווח $[a,b]$ יש בדיוק אותה הסתברות לקרות.	$\frac{1}{b-a+1}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
בינומית $\boldsymbol{Bin(n,p)}$	בחירה של k הצלחות מתוך n ניסויים.	$\binom{n}{k}\cdot p^k\cdot q^{n-k}$	$n\cdot p$	$n\cdot p\cdot q$
גיאומטרית $\boldsymbol{Geom(p)}$	הסתברות ל-$k-1$ כשלונות עד הצלחה בניסוי ה-k.	$q^{k-1}\cdot p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
פואסונית $\boldsymbol{Pois(\lambda)}$	ידוע שאירוע מתרחש בממוצע כל λ יחידות זמן. X סופר כמה אירועים כאלו התרחשו ביחידת זמן אחת.	$\frac{\lambda^k\cdot e^{-\lambda}}{k!}$	λ	λ
בינומית שלילית $\boldsymbol{NB(r,p)}$	ההסתברות ל-k הצלחות עד שקיבלנו r כשלונות.	$\binom{r+k-1}{k}\cdot q^r\cdot p^k$	$\frac{p\cdot r}{q}$	$\frac{p\cdot r}{q^2}$
היפר גיאומטרית $\boldsymbol{HG(N,D,n)}$	יש לנו N כדורים בכד, מתוכם D לבנים,והשאר לא לבנים. מוציאים מהכד n כדורים. x סופר כמה מתוך אלו שהוצאנו היו לבנים.	$\frac{\binom{D}{k}\cdot\binom{N-D}{n-k}}{\binom{N}{n}}$	$\frac{n\cdot D}{N}$	$\frac{n\cdot\left(\frac{D}{N}\right)\cdot\left(1-\frac{D}{N}\right)\cdot(N-n)}{(N-1)}$

תכונות התוחלת:

- $E(X)=\sum_k P(X=k)\cdot k$
- $E(f(X))=\sum_{i=1}^n f(x_i)\cdot p_i$
- $E(c)=c$
- $E(E(X))=E(X)$
- $E(X)=\sum_{i=1}^n P(X=x_i)\cdot x_i=\sum_{i=1}^n x_i\cdot p_i$
- $E(aX+b)=aE(X)+b$
- אם X ו-Y ב"ת $\Leftarrow E(X\cdot Y)=E(X)\cdot E(Y)$
- $E(X+Y|Y=y)=E(X|Y=y)+y$
- $E(X\cdot Y|Y=y)=y\cdot E(X|Y=y)$

תכונות השונות:

- $Var(X)=E\left[\left(X-E(X)\right)^2\right]=E(X^2)-E^2(X)$
- $0\le Var(X)$
- $Var(aX+b)=a^2\cdot Var(X)$
- $Var(c)=0$
- אם X ו-Y תלויים אז:
  $Var(X+Y)=Var(X)+cov(X,Y)+Var(Y)$
- אם X ו-Y ב"ת אז: $Var(X+Y)=Var(X)+Var(Y)$
- סכום משתנים ב"ת: $Var(\sum_{i=1}^n X_i)=\sum_{i=1}^n Var(X_i)$
- סכום משתנים תלויים:
  $Var(\sum_{i=1}^n X_i)=\sum_{i=1}^n\sum_{j=1}^n Cov(X_i,X_j)$

נוסחת בייס: $P(A|B)=\frac{P(B|A)\,P(A)}{P(B)}$

נוסחת בייס כללי: $P(A_k|B)=\frac{P(A_k)\cdot P(B|A_k)}{P(A_1)\cdot P(B|A_1)+\cdots+P(A_n)\cdot P(B|A_n)}$

סטיית תקן: $\sigma(X)=\sqrt{Var(X)}$
$\sigma(aX+b)=|a|\cdot\sigma(X)$

מסקנות: אם חישבנו תוחלת בסעיף כלשהו, ואנחנו צריכים לחשב תוחלת מותנה בסעיף אחר, צריך לחשוב על ההסתברות השלמה וחלוקה למאורעות זרים כדרך למצוא את התוחלת המותנה.
$E(X)=P(A)\cdot E(X|A)+P(A^c)\cdot E(X|A^c)$.

טריק: לפעמים כשיש מספר סופי של הוצאות\ניסויים שאפשר לעשות כדאי לעשות אותם "עד הסוף", לדוגמא אפשר להוציא מספרים מכד עד שנגמר וסכום התוצאות יהיה מספר קבוע כלשהו. כלומר השונות של סכום התוצאות תהיה 0. טוב בשביל לחשב Cov.

כלל השרשרת: $P(\bigcap_{i=1}^n A_i)=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cap A_2)\cdot\ldots\cdot P(A_n|\bigcap_{i=1}^{n-1}A_i)$

נוסחת ההסתברות השלמה: $P(A)=P(B_1)\cdot P(A|B_1)+\cdots+P(B_n)\cdot P(A|B_n)$
נשתמש בנוסחה זו כאשר החישוב הישיר של P(A) מסובך, ועדיף לחלק למקרים.

ההסתברות שA יקרה לפני B:
$P(A|B)=\frac{P(A)}{P(A\cup B)}=\frac{P(A\cap B)}{P(B)}$

נוסחה לחישוב התפלגות סכום מ"מ:

- $P(X+Y=m)=\sum_{k+l=m}P(X=k,Y=l)$
- $P(X+Y=m)=\sum_k P(X=k,Y=m-k)$

דוגמה לחישוב תוחלת מותנה:

מבצעים סדרה של n הטלות מטבע הוגן. יהיה X מספר הפעמים שיצא עץ. לאחר מכן מבצעים סדרה של X הטלות מטבע עם הסתברות $\frac{1}{n}$ להוצאת עץ. יהי Y מספר הפעמים שיצא עץ בסדרת ההטלות השנייה. מהי התוחלת של Y?

$$X\sim Bin\left(n,\frac{1}{2}\right)\Rightarrow E(X)=\frac{n}{2}$$

$$Y|X\sim Bin\left(X,\frac{1}{n}\right)\Rightarrow E(Y|X)=\frac{X}{n}$$

$$E(Y)=E(E(Y|X))=E\left(\frac{X}{n}\right)=\frac{1}{n}E(X)=\frac{1}{n}\cdot\frac{n}{2}=\frac{1}{2}$$

דוגמה לשימוש באינדקטורים:

פקיד צריך לשלוח n מכתבים ל- n יעדים. ברשותו של הפקיד n מעטפות ועליהן כתובים היעדים, כשכל מעטפה מתאימה ליעד אחד בדיוק. הפקיד מחלק את המכתבים למעטפות בצורה מקרית, כאשר ההתפלגות של המעטפה הנבחרת בכל שלב היא אחידה. יהא X מספר המעטפות שהגיעו ליעד. מה התוחלת של X?
נסמן X_i בתור האינדיקטור המתאים למאורע שמכתב i הגיע ליעד. אזי בגלל שההתפלגות אחידה, לכל מכתב ההסתברות שהוא הגיע ליעד היא $\frac{1}{n}$ נשים לב כי: $X=\sum_{i=1}^n X_i$, לפיכך לפי ליניאריות התוחלת:

$$E(X)=\sum_{i=1}^n E(X_i)=\sum_{i=1}^n\frac{1}{n}=\frac{1}{n}\cdot n=1$$

2 משתנים מקריים:

X\Y	Y_1	Y_2	..	Y_m	
X_1	P(1,1)	P(1,2)	..	P(1,m)	$\sum_{j=1}^m P(1,j)$
X_2	P(2,1)	P(2,2)	..	P(2,m)	$\sum_{j=1}^m P(2,j)$
..	..	..	..	..	..
X_n	P(n,1)	P(n,2)	..	P(n,m)	$\sum_{j=1}^m P(n,j)$
	$\sum_{i=1}^n P(i,1)$	$\sum_{i=1}^n P(i,2)$	..	$\sum_{i=1}^n P(i,m)$	1

- התפלגות משותפת: יהיו X,Y מ"מ אזי ההתפלגות המשותפת שלהם: $P(X=k,Y=l)$ (חיתוך של שני המאורעות)
- התפלגות מותנית: חוק ההתפלגות המותנית מקיים ש: $P(X=a|Y=b)=\frac{P(X=a,Y=b)}{P(Y=b)}$ (הערכים בתאים, חלקי השולית)
- X,Y בלתי תלויים אם **לכל** a,b מתקיים: $P(X=a,Y=b)=P(X=a)\cdot P(Y=b)$
- אם מופיע 0 בחוק התפלגות המשותפת (0 בטבלה) אזי המ"מ תלויים (אם לא מופיע אז עדיין צריך לבדוק).
- כדי להראות ש- X,Y תלויים מספיק למצוא a,b כלשהם המקיימים:
  $P(X=a,Y=b)\ne P(X=a)\cdot P(Y=b)$

שונות משותפת: $Cov(X,Y)=E(X\cdot Y)-E(X)E(Y)=E[(X-E(X))(Y-E(Y))]$

- סימטריות: $Cov(X,Y)=Cov(Y,X)$
- $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$
- $Cov(aX,Y)=a\cdot Cov(X,Y)$
- $Cov(X,a)=0$
- $Cov(X,X)=Var(X)$
- X ו-Y ב"ת $\Leftarrow E(X\cdot Y)=E(X)\cdot E(Y)$ ומכאן $Cov(X,Y)=0$, **ההיפך לא בהכרח נכון!**
- $Cov(X,Y)\ne 0 \Leftarrow$ X ו-Y תלויים
- $Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$
- $Var(\sum_{i=1}^n X_i)=\sum_{i=1}^n Var(X_i)+2\cdot\sum_{1\le i<j\le n}Cov(X_i,X_j)$
- $Cov(\sum_{i=1}^n X_i,\sum_{j=1}^m X_j)=\sum_{i=1}^n\sum_{j=1}^m Cov(X_i,X_j)$
- $Cov(X,Y)>0 \Leftarrow$ יש תלות חיובית בין X ל-Y. אם X התרחש אז גדל הסיכוי ש-Y התרחש ולהיפך.
- $Cov(X,Y)<0 \Leftarrow$ יש תלות שלילית בין X ל-Y. אם X התרחש אז קטן הסיכוי ש-Y התרחש ולהיפך.

משפט התוחלת השלמה: לכל Y: $E(X)=E(E(X|Y))=\sum_k P(Y=k)\cdot E(X|Y=k)$

משפט השונות שלמה: $V(X|Y=k)=E(X^2|Y=k)-[E(X|Y=k)]^2$.

מקדם המתאם: $\rho(X,Y)=\frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}=\frac{E(X\cdot Y)-E(X)\cdot E(Y)}{\sqrt{Var(X)\cdot Var(Y)}}$ $(-1\le\rho\le 1)$

- סימטריות: $\rho(X,Y)=\rho(Y,X)$
- $\rho(X,X)=1$
- $\rho(X,Y)=1\Leftrightarrow a>0,\ Y=aX+b$
- $\rho(X,Y)=-1\Leftrightarrow a<0,\ Y=aX+b$
- $\rho(aX+b,cY+d)=\rho(X,Y)\Leftarrow a,c>0$
- $\rho(aX+b,Y)=\frac{a}{|a|}\cdot\rho(X,Y)$
- X ו-Y בלתי מתואמים $\Leftrightarrow Cov(X,Y)=0\Leftrightarrow\rho(X,Y)=0$
- X ו-Y ב"ת $\Leftarrow Cov(X,Y)=0\Leftarrow\rho(X,Y)=0\Leftarrow$ X ו-Y בלתי מתואמים

אי שוויון מרקוב: $P(X\ge a)\le\frac{E(X)}{a}$ כאשר: $(a,X>0)$

שיפור אי-שוויון מרקוב: אם קיים X מ"מ אי-שלילי וידוע שגם X-k (מספר קבוע k) אי-שלילי, עדיף לעשות הזזה של המשתנה למשתנה X-k ואז להשתמש במרקוב.

אי שוויון צ'בישב: $P(|X-E(X)|\ge a)\le\frac{Var(X)}{a^2}$ כאשר $(a>0)$

אינדיקטורים:

משתנה אינדיקטור, הוא משתנה מקרי שמציין האם מאורע מסוים התרחש. כאשר המאורע התרחש, הערך שלו הוא 1, וכאשר המאורע לא התרחש ערכו 0.
שימושי למשל כאשר מעוניינים לחשב תוחלת של משתנה מקרי מסוים שניתן לבטאו כסכום של משתנים מציינים, ואז ניתן להיעזר בליניאריות התוחלת ולהגיע בקלות לתוחלת משתנה זה, אם ידועה ההסתברות להתרחשות המאורעות שהמשתנים המציינים מייצגים.

$X_i=\{^{1,\ \ p}_{0,\ \ 1-p}$ **או** $Bin(1,P(A))\sim X_A=\begin{cases}1, & A\text{ התרחש}\\ 0, & \text{אחרת}\end{cases}$

$E(X_A\cdot X_B)=X_{A\cap B}$, $X_i^k=X_i$, $Var(X_i)=p(1-p)=p-p^2$ $E(X_i)=p$

$Cov(X_i,X_j)=E(X_i\cdot X_j)-E(X_i)E(X_j)=P(X_i=1,X_j=1)-P(X_i=1)\cdot P(X_j=1)$

$X_A\cdot X_B=\begin{cases}1, & A\cap B\text{ התרחש}\\ 0, & \text{אחרת}\end{cases}=X_{A\cap B}$, $A,B\subseteq\Omega$

סכומי סדרות:

- $\sum_{k=1}^n k=\frac{n(n+1)}{2}$
- $\sum_{k=1}^n k^2=\frac{n(n+1)(2n+1)}{6}$
- $\sum_{k=0}^n a\cdot\left(\frac{n}{m}\right)^k=\frac{a}{1-\left(\frac{n}{m}\right)}$
- $\sum_{k=0}^n k\cdot\left(\frac{n}{m}\right)^k=\frac{a}{\left(1-\left(\frac{n}{m}\right)\right)^k}$
- $S_n=\frac{n(a_1+a_2)}{2}$

טורים וזהויות:

- $\sum_{k=0}^n\binom{n}{k}\binom{n}{k}=\binom{2n}{k}$
- $\sum_{k=1}^n k\cdot\binom{n}{k}=n\cdot 2^{n-1}$
- $\sum_{k=0}^n\binom{n}{k}=2^n$
- $\sum_{j=n}^m\binom{j}{n}=\binom{m+1}{n+1}$
- $\sum_{k\,(even)}^n\binom{n}{k}=\sum_{k\,(odd)}^n\binom{n}{k}$
- $P(\bigcup_{i=1}^n A_i)\le\sum_{i=1}^n P(A_i)$
- $\sum_{k=1}^n\binom{n}{k}a^{n-k}b^k=(a+b)^n$
- $\sum_{k=0}^\infty\frac{x^k}{k!}=e^x$
- $\binom{n}{m}=\binom{n-1}{m-1}+\binom{n-1}{m}$

מקבץ שאלות מהחוברת אגודה

יהי $0<p<1$ ויהי $X\sim Geom(p)$ משתנה מקרי בעל התפלגות גיאומטרית.

א. (8 נקודות) הוכיחו שלכל שלם אי שלילי k מתקיים: $P(x>k)=(1-p)^k$.

ב. (10 נקודות)הוכיחו שלכל שלם אי שלילי n ושלם חיובי k מתקיים: $P(X=n+k|X>n)=P(X=k)$.

ג. (7 נקודות) חשבו את התוחלת והשונות של $Y:=\frac{pX-1}{\sqrt{1-p}}$.

א. יהי k שלם אי שלילי. נשים לב שההסתברות ש X גדול מ k שווה להסתברות ש $X=k+1$ או $X=k+2$ או... (עד אינסוף) ולכן: $P(X>k)=\Sigma_{i=k+1}^{\infty}P(X=i)$. (העדפנו לעבור לשוויון כי את השוויון אנו יודעים לפתוח לפי ההתפלגות של X). כעת, נתון כי X מתפלג גאומטרית עם הסתברות p ולכן: $P(X=i)=(1-p)^{i-1}\cdot p$. נציב בסכום ונקבל: $P(X>k)=\Sigma_{i=k+1}^{\infty}(1-p)^{i}\cdot p$.
p אינו תלוי ב i ולכן ניתן להוציאו מחוץ לסכום: $P(X>k)=p\cdot\Sigma_{i=k+1}^{\infty}(1-p)^{i-1}$.
כעת נשים לב שזה סכום סדרה הנדסית אינסופית עם איבר ראשון: $(1-p)^k$ ו- $q=(1-p)$ (מנה קטנה מ 1 כי $0<p<1$) ולכן סכומה הוא: $S=\frac{a_1}{1-q}=\frac{(1-p)^k}{1-(1-p)}=\frac{(1-p)^k}{p}$.
ולכן: $P(X>k)=p\cdot\frac{(1-p)^k}{p}=(1-p)^k$ מש"ל.

ב. נשתמש בנוסחת ההסתברות המותנית: $P(X=n+k|X>n)=\frac{P(X=n+k,X>n)}{P(X>n)}$.
עבור המונה, ניתן לומר ש $P(X=n+k,X>n)=P(X=n+k)$ כי $k>0$ ולכן ההסתברות להיות שווה ל $n+k$ וגם להיות גדול מ n שווה פשוט להסתברות להיות שווה ל $n+k$ כי אז ממילא X יהיה גדול מ n. מכיוון ש X מתפלג גיאומטרית נובע ש: $P(X=n+k)=(1-p)^{n+k-1}\cdot p$.
עבור המכנה, ניתן לומר לפי סעיף א' ש: $P(X>n)=(1-p)^n$.
נציב ונקבל: $P(X=n+k|X>n)=\frac{P(X=n+k,X>n)}{P(X>n)}=\frac{(1-p)^{n+k-1}\cdot p}{(1-p)^n}=(1-p)^{k-1}\cdot p$. ומכאן נשים לב ש $(1-p)^{k-1}\cdot p=P(X=k)$ כי X מתפלג גיאומטרית, מש"ל.

ג. אנו רואים כי Y מוגדר לפי X וקבועים ולכן נחשב את התוחלת והשונות שלו לפי תכונות של תוחלת ושונות.
X מתפלג גיאומטרית ולכן: $E(X)=\frac{1}{p}$, $Var(X)=\frac{1-p}{p^2}$
נשים לב ש p קבוע (לא משתנה מקרי) ומכאן, לפי תכונות של תוחלת:
$E(Y)=E\left(\frac{pX-1}{\sqrt{1-p}}\right)=E\left(\frac{p}{\sqrt{1-p}}X-\frac{1}{\sqrt{1-p}}\right)=\frac{p}{\sqrt{+1-p}}E(X)-\frac{1}{\sqrt{1-p}}=0$.
כעת נחשב את השונות: $Var(Y)=Var\left(\frac{pX-1}{\sqrt{1-p}}\right)=Var\left(\frac{p}{\sqrt{1-p}}X-\frac{1}{\sqrt{1-p}}\right)=\frac{p^2}{1-p}Var(X)=1$ לפי תכונות של שונות.

יהי $n\geq 2$ מספר טבעי. לכל $1\leq i\leq n$ מטילים קוביה הוגנת, כאשר כל הטלות הקוביה בלתי תלויות. יהי X מספר ההטלות עד הפעם הראשונה שהתקבלה התוצאה 3 (כולל ההטלה הנ"ל). אם לא יוצא 3 באף אחת מ- n ההטלות אז נאמר ש- $X=0$. יהי Y משתנה מקרי הסופר את מספר הפעמים שהתקבל הרצף 66, כלומר את מספר האינדקסים $1\leq i\leq n-1$ כך שתוצאות ההטלות ה- i וה- $i+1$ הן 6.

א. (8 נקודות) חשבו את ההתפלגות של X.
ב. (11 נקודות) חשבו את התוחלת של Y בהינתן המאורע $X=0$.
ג. (6 נקודות) האם X ו- Y בלתי תלויים? נמקו את תשובתכם.

א. נשים לב שטווח הערכים של X הוא $\{0,1,\dots,n\}$ (לכל היותר n הטלות או שלא יצא בכלל 3 ואז $X=0$) אנו צריכים לחשב את $P(X=k)$ לכל k בטווח.
עבור $k=0$: $P(X=0)=\left(\frac{5}{6}\right)^n$ כי דורשים שלא יצא 3 באף הטלה אז מותר שיצאו בכל אחת מ n ההטלות רק 5 מתוך 6 המספרים שעל הקובייה.
עבור $k\neq 0$: $P(X=k)=\left(\frac{5}{6}\right)^{k-1}\cdot\frac{1}{6}$ כי דורשים שב $k-1$ ההטלות הראשונות לא יצא 3 ובהטלה ה k כן יצא 3.

ב. יש כאן ספירה מורכבת (תוצאה של 2 אינדקסים) ולכן נשתמש באינדיקטורים: לכל $1\leq i\leq n-1$ נגדיר את Y_i להיות האינדיקטור השווה ל 1 אם ורק אם בהטלות $(i,i+1)$ יצא $(6,6)$. אז: $Y=\Sigma_{i=1}^{n-1}Y_i$.
כעת, נחשב את התוחלת של $Y|X=0$ לפי תכונת ליניאריות התוחלת:
$E(Y|X=0)=E(\Sigma_{i=1}^{n-1}Y_i|X=0)=\Sigma_{i=1}^{n-1}E(Y_i|X=0)$. (יש לשים לב שאנו מתייחסים לכל הביטוי: $Y|X=0$ כמשתנה מקרי חדש, ובאותו אופן: $Y_i|X=0$)
נותר לחשב את $E(Y_i|X=0)$ לכל $1\leq i\leq n-1$.
נזכיר כי $Y_i|X=0$ הוא אינדיקטור (כי עדיין מקבל ערכים כמו Y_i) ולכן:
$E(Y_i|X=0)=\mathrm{P}(\mathrm{Y_i}=1|\mathrm{X}=0)=\frac{1}{5}\cdot\frac{1}{5}=\frac{1}{25}$ כי אנו דורשים שיצא 6 ב 2 ההטלות $(i,i+1)$
ומכיוון שידוע ש $X=0$ אז 3 לא יכול לצאת כתוצאה ולכן נותרנו עם 5 מספרים ואנו דורשים שיצא 6. ולכן ההסתברות לכל הטלה היא $\frac{1}{5}$.

ג. המשתנים כן תלויים. נראה דוגמא:
$P(X=1,Y=n-1)=0$ כי אם $X=1$ זה אומר שיצא 3 בפעם הראשונה. אבל: $Y=n-1$ וזה אומר שיצאו 6 בכל ההטלות – מה שאינו הגיוני. מצד שני:
$P(X=1)=\frac{1}{6}>0$ וגם $P(Y=n-1)=\left(\frac{1}{6}\right)^n>0$ ולכן המכפלה ביניהם תהיה שונה מ 0 ולכן לא שווה ל $P(X=1,Y=n-1)$.

לאריאל יש שלושה זוגות מכנסיים בצבעים שיסומנו ב- 1,2,3 וארבע חולצות בצבעים שיסומנו ב - 1,2,3,4. במשך 90 ימים, בכל יום הוא בוחר מכנסיים וחולצה ללבוש באופן הבא: תחילה הוא בוחר מכנסיים באופן מקרי אחיד (מבין שלושה הזוגות שברשותו) ואז הוא בוחר חולצה באופן אחיד מבין החולצות שברשותו שצבען שונה מצבע המכנסיים שבחר. הבחירות בימים שונים בלתי תלויות.

א. (9 נקודות) יהי X מספר המכנס הנבחר ביום כלשהו ויהי Y מספר החולצה שנבחרה באותו יום. חשבו את ההתפלגות המשותפת של X ו- Y.
ב. (6 נקודות) חשבו את $P(X=1|Y=3)$.
ג. (10 נקודות) יהי Z משתנה מקרי הסופר את מספר הימים (מתוך ה-90) בהם אריאל לבש את חולצה מספר 4. הוכיחו ש- $P(|Z-30|\geq 10)\leq\frac{1}{5}$.

א. טווח הערכים של X הוא $\{1,2,3\}$, טווח הערכים של Y הוא $\{1,2,3,4\}$. נחשב תחילה את כל החיתוכים של ההסתברויות לערכי X,Y:
נשים לב שלא ייתכן ש $X=Y$ כי אריאל בוחר חולצה בצבע השונה מהמכנס שבחר ולכן:
$P(X=k,Y=k)=0$ לכל $1\leq k\leq 3$.
אחרת, לפי נוסחת הסתברות מותנית (כי קל לחשב כאן – ברור למה Y תלוי ב X)
$P(X=k,Y=m)=P(X=k)\cdot P(Y=m|X=k)=\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9}$ לכל $1\leq k\leq 3$ ולכל $1\leq m\leq 4$ פרט ל $k=m$. כי ההסתברות לבחור מכנס (בצבע ספציפי) היא 1/3 (אחידה) ואז בהינתן שמכנס k נבחר, ההסתברות לבחור את החולצה בצבע m (כאשר $m\neq k$) היא 1 מתוך ה 3 אפשריות (מורידים את החולצה שהיא באותו הצבע של הכנס שנבחר כי היא לא יכולה להיבחר).
תשובה סופית: (אפשרי לעשות גם כאן טבלה אבל מכיוון שמצאנו באופן כללי אז נשתמש בנוסחא שמצאנו)
$P(X=k,Y=m)=\begin{cases}\frac{1}{9}, \text{if } m\neq k\\ 0, \text{if } m=k\end{cases}$.

ב. לפי נוסחת ההסתברות המותנית: $P(X=1|Y=3)=\frac{P(X=1,Y=3)}{P(Y=3)}$. לפי א' $P(X=1,Y=3)=\frac{1}{9}$ ולפי סכימת ערכים (בטבלה זו סכימת שורה או עמודה) נחשב את $P(Y=3)$ ונקבל:
$P(Y=3)=P(X=1,Y=3)+P(X=2,Y=3)+P(X=3,Y=3)=\frac{1}{9}+\frac{1}{0}+0=\frac{2}{9}$.
נציב הכל ונקבל ש: $P(X=1|Y=3)=\frac{P(X=1,Y=3)}{P(Y=3)}=\frac{\frac{1}{9}}{\frac{2}{9}}=\frac{1}{2}$.

ג. נשים לב ש Z סופר הצלחות (בחירת חולצה 4) מתוך כמות נסיונות ידועה (90 יום) כאשר כל הניסויים בלתי תלויים זה בזה ויש לכל ניסוי סיכוי להצלחה ולכישלון ולכן: $Z\sim Bin(90,p)$ כאשר לפי א':
$p=P(Y=4)=P(X=1,Y=3)+P(X=2,Y=3)+P(X=3,Y=3)=\frac{1}{9}+\frac{1}{9}+\frac{1}{9}=\frac{1}{3}$.
ולכן: $E(Z)=90\cdot\frac{1}{3}=30$, $Var(Z)=90\cdot\frac{1}{3}\cdot\frac{2}{3}=20$ לפי נוסחאות לתוחלת ושונות של התפלגות בינומית.
נשים לב שהביטוי נראה כמו אי שיוויון צ'בישב ולכן נשתמש באי שוויון זה:
$P(|Z-30|\geq 10)=P(|Z-E(Z)|\geq 10)\leq\frac{Var(Z)}{10^2}=\frac{20}{100}=15$ משל.

נתון מטבע הנותן 1 בהסתברות $1/3$ ו- 0 בהסתברות $2/3$. מטילים את המטבע 4 פעמים כאשר ההטלות בלתי תלויות. יהי X סכום תוצאות 2 ההטלות הראשונות, יהי Y סכום תוצאת ההטלה השנייה והשלישית ויהי Z סכום תוצאות שתי ההטלות האחרונות.

א. (15 נקודות) חשבו את ההתפלגות המשותפת של X ו $-Y$.
ב. (5 נקודות) האם X ו- Y בלתי תלויים? הוכיחו את תשובתכם.
ג. (7 נקודות) חשבו את $P(X+Y+Z=4)$.
ד. (7 נקודות) חשבו את $P(X=1,Y=1|Z=1)$.

א. עבור X,Y,Z טווח הערכים שהם מקבלים הוא: $\{0,1,2\}$ - סופי ולכן נשתמש בטבלה:
נחשב את כל החיתוכים של ערכי X וערכי Y:
$P(X=0,Y=0)=\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}=\frac{8}{27}$ – נדרוש שב 3 ההטלות הראשונות יצא 0 במטבע.
$P(X=1,Y=0)=\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}=\frac{4}{27}$ – נדרוש בהטלה ה 1 שיצא אחד ובשתיים האחרות 0.
$P(X=2,Y=0)=\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}=0$ – אם $X=2$ אז בהטלה ה 2 יצא 1 ואז Y לא יכול להיות 0.
$P(X=0,Y=1)=\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}=\frac{4}{27}$ – נדרוש שב 2 ההטלות הראשונות יצא 0 ובשלישית 1.
$P(X=1,Y=1)=\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}+\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{6}{27}$ – נדרוש (1,0,1) או (0,1,0) כי ההטלה ה 2 נספרת גם ל X וגם ל Y.
$P(X=2,Y=1)=\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{2}{27}$ – נדרוש ב 2 ההטלות הראשונות 1 ובשלישית 0.
$P(X=0,Y=2)=0$ – אם $Y=2$ אז בהטלה ה 2 יצא 1 ואז X לא יכול להיות 0.
$P(X=1,Y=2)=\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}=\frac{2}{27}$ – נדרוש בהטלה ה 1 שיצא 0 ובשתיים האחרות 1.
$P(X=2,Y=2)=\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{27}$ – נדרוש 1 ב 3 ההטלות.
תשובה סופית:

X / Y	0	1	2	Σ
0	$\frac{8}{27}$	$\frac{4}{27}$	0	$\frac{12}{27}$
1	$\frac{4}{27}$	$\frac{6}{27}$	$\frac{2}{27}$	$\frac{12}{27}$
2	0	$\frac{2}{27}$	$\frac{1}{27}$	$\frac{3}{27}$
Σ	$\frac{12}{27}$	$\frac{12}{27}$	$\frac{3}{27}$	1

ב. המשתנים תלויים. דוגמא: $P(X=2,Y=0)=0$ לפי הטבלה בסעיף א'.
מצד שני: $P(X=2)=0+\frac{2}{27}+\frac{1}{27}=\frac{3}{27}$ (סכימת עמודה), $P(Y=0)=\frac{8}{27}+\frac{4}{27}+0=\frac{12}{27}$ (סכימת שורה), ורואים ש $P(X=2)\cdot P(Y=0)=\frac{3}{27}\cdot\frac{12}{27}\neq 0$ ולכן המשתנים תלויים.

ג. כדי לחשב את $P(X+Y+Z=4)$ נשים לב שעבור הסדרה (a,b,c,d) של הטלות המטבע. אם יש 1 ב b הוא נספר פעמיים: פעם אחת ב X ופעם אחת ב Y, כמו כן, אם יש 1 ב c אז הוא נספר פעמיים: פעם אחת ב Y ופעם אחת ב Z. עבור 1 ב a או ב d הוא נספר פעם אחת. ולכן הסדרות שייתנו 4 בס"ה הן: $\{(0,1,1,0),(1,1,0,1),(1,0,1,1)\}$ ומכאן:
$P(X+Y+Z=4)=\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}+\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}=\frac{4}{3^4}+\frac{2}{3^4}+\frac{2}{3^4}=\frac{8}{81}$.

ד. לפי נוסחת ההסתברות המותנה: $P(X=1,Y=1|Z=1)=\frac{P(X=1,Y=1,Z=1)}{P(Z=1)}$. נחשב את ההסתברויות:
$P(X=1,Y=1,Z=1)=P(\{(0,1,0,1),(1,0,1,0)\})=\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{8}{81}$ כי צריך שהסכום יהיה 3 ובאותו האופן כמו בסעיף הקודם.
עבור $P(Z=1)$ נדרוש שבהטלה ה 3 יצא 1 וברביעית 0 או להיפך ולכן: $P(Z=1)=2\cdot\frac{2}{3}\cdot\frac{1}{3}=\frac{4}{9}$.
נציב ונקבל: $P(X=1,Y=1|Z=1)=\frac{P(X=1,Y=1,Z=1)}{P(Z=1)}=\frac{\frac{8}{81}}{\frac{4}{9}}=\frac{2}{9}$.

בכל יום במהלך חודש בן 30 יום אריאל אוכל פרי אחד. על מנת להחליט איזה פרי לאכול, הוא מטיל קובייה הוגנת כאשר הטלות קובייה בימים שונים הן בלתי תלויות. אם תוצאת הקובייה היא 1,2 או 3 הוא אוכל תפוח, אם היא 4 או 5 הוא אוכל אפרסק ואם היא 6 הוא אוכל בננה. יהי X מספר התפוחים הכולל שאריאל אכל, יהי Y מספר האפרסקים הכולל שאריאל אכל ויהי Z מספר הבננות הכולל שאריאל אכל.

א. (6 נקודות) חשבו את תוחלת ושונות X.
ב. (15 נקודות) חשבו את $\rho(Y,Z)$.
ג. (12 נקודות) הוכיחו ש- $P(X+Z\geq 25)\leq\frac{4}{15}$.

א. נשים לב שיש כאן 30 ניסויים (ימים בהם אריאל מטיל קוביה) כאשר הניסויים בלתי תלויים ובכל ניסוי יש סיכוי להצלחה וכשלון. X סופר את ההצלחות (הצלחה במובן שיצא 1,2,3 ואריאל אכל תפוח), Y סופר את מספר ההצלחות (במובן שיצא 4,5 ואריאל אכל אפרסק), Z סופר את מספר ההצלחות (במובן שיצא 6 ואריאל אכל בננה). לכן כל המשתנים מתפלגים בינומית:
$X\sim Bin\left(30,\frac{1}{2}\right), Y\sim Bin\left(30,\frac{1}{3}\right), Z\sim Bin\left(30,\frac{1}{6}\right)$.
לכן, לפי הנוסחאות לתוחלת ושונות התפלגות בינומית:
$E(X)=30\cdot\frac{1}{2}=15$, $Var(X)=30\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{15}{2}$.

ב. כדי לחשב את מקדם המתאם, עלינו תחילה לחשב את $Cov(Y,Z)$. לכן, כדי להקל על החישוב, נפרק את Y,Z לאינדיקטורים. נגדיר לכל $1\leq i\leq 30$ את Y_i – שווה 1 אם ורק אם אריאל אכל אפרסק ביום i.
Z_i - שווה 1 אם ורק אם אריאל אכל בננה ביום i. מכאן: $Y=\Sigma_{i=1}^{30}Y_i$, $Z=\Sigma_{i=1}^{30}Z_i$.
לפי תכונת שונות משותפת לסכום מול סכום:
$Cov(Y,Z)=Cov(\Sigma_{i=1}^{30}Y_i,\Sigma_{i=1}^{30}Z_i)=\Sigma_{i=1}^{30}\Sigma_{j=1}^{30}Cov(Y_i,Z_j)$, נותר לחשב את $Cov(Y_i,Z_j)$ לכל i,j:
נשים לב שאם $i\neq j$ אז Y_i,Z_j בלתי תלויים כי ההטלות בימים שונים הן בלתי תלויות ולכן:
$Cov(Y_i,Z_j)=0$. נותר לחשב Cov רק עבור $i=j$:
לפי הנוסחא לחישוב ישיר של השונות המשותפת:
$Cov(Y_i,Z_i)=E(Y_iZ_i)-E(Y_i)\cdot E(Z_i)$, כל המשתנים הם אינדיקטורים ולכן:
$E(Y_iZ_i)=P(Y_iZ_i=1)=P(Y_i=1,Z_i=1)=0$ כי לא ייתכן שאריאל אכל גם אפרסק וגם בננה כי הוא מטיל את הקוביה רק פעם אחת ועבור כל תוצאה הוא אוכל פרי אחד.
$E(Y_i)=\frac{1}{3}$ כי עבור התוצאות 4,5 אריאל אוכל אפרסק.
$E(Z_i)=\frac{1}{6}$ כי רק אם יוצא 6 אריאל אוכל בננה.
לכן: $Cov(Y_i,Z_i)=0-\frac{1}{3}\cdot\frac{1}{6}=-\frac{1}{18}$.
נחזור ונציב למעלה:
$Cov(Y,Z)=\Sigma_{i=1}^{30}\Sigma_{j=1}^{30}Cov(Y_i,Z_j)=\Sigma_{i=1}^{30}Cov(Y_i,Z_i)=\Sigma_{i=1}^{30}\left(-\frac{1}{18}\right)=-\frac{30}{18}$.
נמשיך ונחשב את $Var(Y)$: לפי תכונה של שונות של סכום משתנים בלתי תלויים (כי ההטלות בימים שונים בלתי תלויות: $Var(Y)=Var(\Sigma_{i=1}^{30}Y_i)=\Sigma_{i=1}^{30}Var(Y_i)=\Sigma_{i=1}^{30}\left(\frac{1}{3}\right)\cdot\left(\frac{2}{3}\right)=\frac{60}{9}=\frac{20}{3}$ כאשר החישוב של $Var(Y_i)$ לכל i הוא לפי נוסחת השונות של משתנה ברנולי (אינדיקטור).
באותו אופן עבור Z: $Var(Z)=Var(\Sigma_{i=1}^{30}Z_i)=\Sigma_{i=1}^{30}Var(Z_i)=\Sigma_{i=1}^{30}\left(\frac{1}{6}\right)\cdot\left(\frac{5}{6}\right)=\frac{150}{36}=\frac{25}{6}$.
נציב הכל לנוסחת חישוב מקדם המתאם:
$\rho(Y,Z)=\frac{Cov(Y,Z)}{\sqrt{Var(Y)}\sqrt{Var(Z)}}=-\frac{\frac{30}{18}}{\sqrt{\frac{20}{3}}\sqrt{\frac{25}{6}}}=-\frac{1}{\sqrt{10}}$.

ג. נשים לב שמכיוון ש $X+Y+Z=30$ (כי בכל יום אריאל אוכל פרי (תפוח או אפרסק או בננה) ולכן: $P(X+Z\geq 25)=P(Y\leq 5)$,
נזכיר כי $Y\sim Bin\left(30,\frac{1}{3}\right)$ ולכן: $E(Y)=\frac{30}{3}=10$, $Var(Y)=30\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{20}{3}$
נשתמש באי שוויון צ'בישב (כי הסימן הפוך אז מרקוב לא יעזור): נשתמש בטריקים של העברת אגפים והוספה ל 2 הצדדים כדי שהביטוי בתוך ההסתברות יהיה דומה לביטוי באי השוויון של צ'בישב:
$P(Y\leq 5)=P(-Y\geq -5)=P(-Y+10\geq -5+10)=P(10-Y\geq 5)$, נמשיך ב"טריקים" ונשים לב ש $P(10-Y\geq 5)\leq P(|10-Y|\geq 5)$ כי כשיש ערך מוחלט זה מתחלק לסכום של 2 הסתברויות, אחת שהייתה $(P(10-Y\geq 5))$ + אחת בכיוון השלילי $(P(10-Y\leq -5))$ ולכן $P(|10-Y|\geq 5)$ גדול או שווה ל $P(10-Y\geq 5)$. לכן:
$P(10-Y\geq 5)\leq P(|10-Y|\geq 5)=\mathrm{P}(|\mathrm{Y}-10|\geq 5)=\mathrm{P}(|\mathrm{Y}-\mathrm{E}(\mathrm{Y})|\geq 5)$. הפכנו את ההפרש בתוך הערך המוחלט ומתקיים: $E(Y)=10$. למעשה, השתמשנו בטריקים כדי להגיע לביטוי מהצורה $\mathrm{P}(|\mathrm{Y}-\mathrm{E}(\mathrm{Y})|\geq 5)$ כדי שנוכל להשתמש באי השוויון של צ'בישב ולכן הורדנו 10 מ 2 האגפים כי ידענו ש 10 היא התוחלת של Y. כעת נשתמש באי השוויון ונקבל:
$\mathrm{P}(|\mathrm{Y}-\mathrm{E}(\mathrm{Y})|\geq 5)\leq\frac{\mathrm{Var}(\mathrm{Y})}{5^2}=\frac{\frac{20}{3}}{25}=\frac{20}{75}=\frac{4}{15}$ ולכן: $P(X+Z\geq 25)\leq\frac{4}{15}$ מש"ל.

יהי $N\sim Poi(\lambda)$ משתנה מקרי בעל התפלגות פואסון עם פרמטר $\lambda>0$. מטילים מטבע הוגן N פעמים כאשר ההטלות בלתי תלויות. יהי X משתנה מקרי הסופר את מספר ההטלות שתוצאתו עץ.

א. (11 נקודות) חשבו את $E(X)$.
ב. (6 נקודות) האם X ו- N בלתי תלויים? נמקו את תשובתכם.
ג. (8 נקודות) הוכיחו ש $P(X\geq\lambda)\leq 1/2$.

א. נשים לב ש X תלוי ב N כי אם ידוע לנו כמה N, יהיה קל למנות כמה פעמים יצא עץ. מכאן, נשים לב שכאשר ידוע ש $N=n$, X מתפלג בינומית כי יש מספר נסיונות ידוע, כל נסיון בלתי תלוי באחר, לכל נסיון יש הצלחה או כשלון ומתוכם X סופר את מספר ההצלחות (קבלת עץ). לכן: $X|N=n\sim Bin\left(n,\frac{1}{2}\right)$ (ההסתברות 1/2 היא לקבלת עץ במטבע).
נחשב את $E(X)$ באמצעות נוסחת התוחלת השלמה ולפי הנוסחאות להתפלגויות השונות:
$E(X)=E(E(X|N))=\Sigma_{n=0}^{\infty}P(N=n)\cdot E(X|N=n)=\Sigma_{n=1}^{\infty}\left(\frac{\lambda^n\cdot e^{-\lambda}}{n!}\right)\cdot\frac{n}{2}$ (ניתן להתחיל את הסכום מ 1 כי עבור $n=0$ מכפילים ב 0), נצמצם את n, נוציא λ אחת ונסדר: $\Sigma_{n=1}^{\infty}\left(\frac{\lambda^{n-1}\cdot e^{-\lambda}}{(n-1)!}\right)\cdot\frac{\lambda}{2}$,
נבצע הצבה: $t=n-1$ ונקבל: $\Sigma_{t=0}^{\infty}\left(\frac{\lambda^t\cdot e^{-\lambda}}{t!}\right)\cdot\frac{\lambda}{2}$. כעת נשים לב שהסכום (פרט ל $\frac{\lambda}{2}$) הוא כל ההתפלגות של משתנה פואסוני (סוכמים את כל הערכים שלו) ולכן סכום זה הוא 1. מכאן: $E(X)=\frac{\lambda}{2}$.

ב. הם כן תלויים. דוגמא: $P(X=2,N=1)=0$ כי לא ייתכן שיצא פעמיים עץ אם הייתה רק הטלה אחת. מצד שני: $P(X=2)>0$ כי יש סיכוי לפעמיים עץ (אם יהיו לפחות 2 הטלות ובשתיהן יצא עץ) וגם $P(N=1)=\lambda\cdot e^{-\lambda}>0$ כי $\lambda>0$. ולכן: $P(X=2)\cdot P(X=1)\neq 0$.

ג. מדובר באי שוויון ולכן נשתמש באי שוויון מרקוב (ובסעיף א'): $P(X\geq\lambda)\leq\frac{E(X)}{\lambda}=\frac{\frac{\lambda}{2}}{\lambda}=\frac{1}{2}$. מש"ל.

יהי $n\geq 3$ מספר טבעי. לכל $1\leq i\leq n$ מטילים מטבע הוגן שצדדיו הם 0 ו-1, כאשר כל הטלות המטבע בלתי תלויות. יהי X משתנה מקרי הסופר את מספר הפעמים שהתקבל הרצף 10, כלומר את מספר האינדקסים $1\leq i\leq n-1$ כך שתוצאת ההטלה ה i היא 1 ותוצאת ההטלה ה- $i+1$ היא 0. יהי Y משתנה מקרי הסופר את מספר הפעמים שהתקבל הרצף 101.

א. (6 נקודות) חשבו את התוחלת של Y.
ב. (9 נקודות) חשבו את השונות של X.
ג. (10 נקודות) חשבו את השונות המשותפת $Cov(X,Y)$ של X ושל Y.

א. מכיוון שהספירה מורכבת נפרק את המשתנים לסכום אינדיקטורים:
נגדיר לכל $1\leq i\leq n-1$ את X_i להיות אינדיקטור השווה ל 1 אם תוצאת ההטלה ה i הייתה 1 וגם תוצאת ההטלה ה $i+1$ הייתה 0. מכאן: $X=\Sigma_{i=1}^{n-1}X_i$.
כמו כן: נגדיר לכל $1\leq i\leq n-2$ את Y_i להיות אינדיקטור השווה ל 1 אם תוצאת ההטלה ה i הייתה 1 וגם תוצאת ההטלה ה $i+1$ הייתה 0 וגם תוצאת ההטלה ה $i+2$ הייתה 1. מכאן: $Y=\Sigma_{i=1}^{n-2}Y_i$.
כעת, נחשב את התוחלת של Y לפי תכונת ליניאריות התוחלת: $E(Y)=E(\Sigma_{i=1}^{n-2}Y_i)=\Sigma_{i=1}^{n-2}E(Y_i)$.
נותר לחשב את $E(Y_i)$ לכל i. מכיוון ש Y_i אינדיקטור אז:
$E(Y_i)=P(Y_i=1)=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{8}$ – נדרוש 101 בדיוק בהטלות: $i,i+1,i+2$.
נציב ונקבל: $E(Y)=\Sigma_{i=1}^{n-2}E(Y_i)=\Sigma_{i=1}^{n-2}\left(\frac{1}{8}\right)=\frac{n-2}{8}$.

ב. לפי תכונות של שונות של סכום משתנים:
$Var(X)=Var(\Sigma_{i=1}^{n-1}X_i)=Var(\Sigma_{i=1}^{n}X_i)=\Sigma_{i=1}^{n-1}Var(X_i)+2\cdot\Sigma_{i<j}Cov(X_i,X_j)$. נעבור לחישוב $Var(X_i)$: לפי נוסחא לשונות של אינדיקטור: $Var(X_i)=P(X_i=1)\cdot P(X_i=0)$.
$P(X_i=1)=\left(\frac{1}{2}\right)^2=\frac{1}{4}$ כי נדרוש 10 בהטלות ה $i,i+1$.
נציב ונקבל: $Var(X_i)=P(X_i=1)\cdot P(X_i=0)=\frac{1}{4}\cdot\frac{3}{4}=\frac{3}{16}$.
כעת נעבור לחישוב $Cov(X_i,X_j)$ לכל $i<j$.
לפי תכונה של שונות משותפת, אם X_i,X_j בלתי תלויים אז $Cov(X_i,X_j)=0$.
היות וההטלות בלתי תלויות זו בזו נובע שהמשתנים יהיו תלויים רק אם הם יתייחסו לאותן הטלות (יהיה חיתוך). דבר זה ייקרה רק אם $j=i+1$ כי אז ההטלה ה $i+1$ משותפת.
מכאן נחשב עבור $j=i+1$ לפי הנוסחא לחישוב שונות משותפת:
$Cov(X_i,X_{i+1})=E(X_iX_{i+1})-E(X_i)\cdot E(X_{i+1})$.
חישבנו את $P(X_i=1)=\frac{1}{4}$ ולכן: $E(X_{i+1})=E(X_i)=\frac{1}{4}$.
נותר לחשב את $E(X_iX_{i+1})$: מכפלת אינדיקטורים הוא אינדיקטור ולכן:
$E(X_iX_{i+1})=P(X_iX_{i+1}=1)=P(X_i=1,X_{i+1}=1)=0$ כי אם בהטלה ה i יצא 1 ובהטלה ה $i+1$ יצא 0 אז כבר לא ייתכן שבהטלה ה $i+1$ יצא 1.
לכן נציב הכל ונקבל: $Cov(X_i,X_{i+1})=0-\frac{1}{4}\cdot\frac{1}{4}=-\frac{1}{16}$.
מכאן:
$Var(X)=\Sigma_{i=1}^{n-1}Var(X_i)+2\cdot\Sigma_{i<j}Cov(X_i,X_j)=\Sigma_{i=1}^{n-1}Var(X_i)+2\cdot\Sigma_{i=1}^{n-2}Cov(X_i,X_{i+1})$,
נציב את מה שקיבלנו ונחשב:
$=\Sigma_{i=1}^{n-1}\left(\frac{3}{16}\right)+2\cdot\Sigma_{i=1}^{n-2}\left(-\frac{1}{16}\right)=\frac{3(n-1)}{16}-\frac{2(n-2)}{16}=\frac{n+1}{16}$.

ג. לפי תכונות של שונות משותפת של סכום מול סכום:
$Cov(X,Y)=Cov(\Sigma_{i=1}^{n-1}X_i,\Sigma_{i=1}^{n-2}Y_i)=\Sigma_{i=1}^{n-1}\Sigma_{j=1}^{n-2}Cov(X_i,Y_j)$. נותר לחשב $Cov(X_i,Y_j)$ לכל i,j.
נשים לב שאם X_i,Y_j אינם תלויים אז $Cov(X_i,Y_j)=0$. לכן נחשב כאשר הם תלויים.
היות וההטלות אינן תלויות, ייתכן ו X_i,Y_j תלויים רק אם הם חולקים הטלות משותפות.
Y_j מתייחס להטלות: $j,j+1,j+2$.
ו- X_i מתייחס להטלות $i,i+1$.
לכן יהיו חיתוכים אם: $j=i+1$, $j=i$, $j=i-1$, $j=i-2$
נחשב: (לפי חישוב של שונות משותפת)
את התוחלות של X_i,Y_i חישבנו לעיל לכל i: $E(X_i)=\frac{1}{4}$, $E(Y_i)=\frac{1}{8}$.
עבור $j=i-2$. $Cov(X_i,Y_{i-2})=E(X_iY_{i-2})-E(X_i)\cdot E(Y_{i-2})$. כאשר:
$E(X_iY_{i-2})=P(X_i=1,Y_{i-2}=1)=\frac{1}{16}$ כאשר השוויון הראשון הוא כי X_iY_{i-2} הוא אינדיקטור ויקבל 1 רק כאשר 2 המשתנים בעצמם 1. והשוויון השני הוא כי אנו דורשים את הרצף (1,0,1,0) במקומות: (i-2,i-1,i,i+1) בהתאמה וישנה הסתברות של 1/2 לכל תוצאה.
ולכן: $Cov(X_i,Y_{i-2})=E(X_iY_{i-2})-E(X_i)\cdot E(Y_{i-2})=\frac{1}{16}-\frac{1}{4}\cdot\frac{1}{8}=\frac{1}{32}$.
עבור $j=i-1$. $Cov(X_i,Y_{i-1})=E(X_iY_{i-1})-E(X_i)\cdot E(Y_{i-1})$. כאשר:
$E(X_iY_{i-1})=P(X_i=1,Y_{i-1}=1)=0$ כי אם $Y_{i-1}=1$ אז היה הרצף (1,0,1) במקומות: (i-1,i,i+1) ואז במקומות (i,i+1) אין את הרצף 10.
ולכן: $Cov(X_i,Y_{i-1})=E(X_iY_{i-1})-E(X_i)\cdot E(Y_{i-1})=0-\frac{1}{4}\cdot\frac{1}{8}=-\frac{1}{32}$.
עבור $j=i$. $Cov(X_i,Y_{i-1})=E(X_iY_i)-E(X_i)\cdot E(Y_i)$. כאשר:
$E(X_iY_i)=P(X_i=1,Y_i=1)=\frac{1}{8}$ כי אם $Y_i=1$ אז היה הרצף (1,0,1) במקומות: (i,i+1,i+2) ואז במקומות (i,i+1) כבר יש את הרצף 10. ולכן ההסתברות היא 1/2 לכל מקום ובסה"כ 1/8.
ולכן: $Cov(X_i,Y_{i-1})=E(X_iY_{i-1})-E(X_i)\cdot E(Y_{i-1})=\frac{1}{8}-\frac{1}{4}\cdot\frac{1}{8}=\frac{3}{32}$.
עבור $j=i+1$. $Cov(X_i,Y_{i+1})=E(X_iY_{i+1})-E(X_i)\cdot E(Y_{i+1})$. כאשר:
$E(X_iY_{i+1})=P(X_i=1,Y_{i+1}=1)=0$ כי אם $Y_{i+1}=1$ אז היה הרצף (1,0,1) במקומות: (i+1,i+2,i+3) ואז במקומות (i,i+1) לא יכול להיות הרצף 10. (כי במקום i+1 יש כבר 1)
ולכן: $Cov(X_i,Y_{i+1})=E(X_iY_{i+1})-E(X_i)\cdot E(Y_{i+1})=0-\frac{1}{4}\cdot\frac{1}{8}=-\frac{1}{32}$.
נציב הכל ונקבל:
$Cov(X,Y)=\Sigma_{i=1}^{n-1}\Sigma_{j=1}^{n-2}Cov(X_i,Y_j)=\Sigma_{i=1}^{n-1}Cov(X_i,Y_{i-2})+\Sigma_{i=1}^{n-1}Cov(X_i,Y_{i-1})+\Sigma_{i=1}^{n-2}Cov(X_i,Y_i)+\Sigma_{i=1}^{n-3}Cov(X_i,Y_{i+1})$,
$=\Sigma_{i=1}^{n-1}\left(\frac{1}{32}\right)+\Sigma_{i=1}^{n-1}\left(-\frac{1}{32}\right)+\Sigma_{i=1}^{n-2}\left(\frac{3}{32}\right)+\Sigma_{i=1}^{n-3}\left(-\frac{1}{32}\right)=\frac{n-1}{32}-\frac{n-1}{32}+\frac{3(n-2)}{32}-\frac{n-3}{32}=\frac{2n-3}{32}$
כאשר תחומי הסיגמא נקבעו לפי כמות משתני X_i $(n-1)$ וכמות משתני Y_i $(n-2)$ שנכנסים אבל נדרוש שלא תהיה חריגה.

א. (6 נקודות) יהא X משתנה מקרי המקבל רק ערכים שלמים אי שליליים. הוכיחו ש- $E(X)=\Sigma_{k=0}^{\infty}P(X>k)$. ניתן להשתמש בפעולות שונות על טורים (כגון הפרדה לסכומים, שינוי סדר סכימה וכו') ללא הסבר מדוע זה מותר.

ב. (10 נקודות) מטילים קובייה הוגנת מספר פעמים כאשר כל ההטלות בלתי תלויות. המספרים 1 ו-2 שעל הקובייה צבועים באדום, המספרים 3 ו-4 בכחול והמספרים 5 ו-6 בצהוב. לכל k טבעי חשבו את ההסתברות שב- k ההטלות הראשונות התקבלו מספרים משני צבעים שונים לכל היותר (שימו לב גם לערכי k קטנים).

ג. (9 נקודות) יהי X משתנה מקרי הסופר את מספר ההטלות הנדרשות בניסוי המתואר בסעיף הקודם עד שמתקבלים בפעם הראשונה כל שלושת הצבעים. הוכיחו ש- $E(X)=5.5$ (שוב שימו לב לערכים הקטנים).

א. ננסה לפשט את $\Sigma_{k=0}^{\infty}P(X>k)$ ולהגיע ל – $E(X)$. נשים לב ש: $P(X>k)=\Sigma_{i=k+1}^{\infty}P(X=i)$ כי ההסתברות ש X יהיה גדול מ k שווה להסתברות ש X יהיה שווה ל $k+1$ או ש X יהיה שווה ל $k+2$ או... עד אינסוף. ולכן נוכל להציב ולקבל: $\Sigma_{k=0}^{\infty}P(X>k)=\Sigma_{k=0}^{\infty}\Sigma_{i=k+1}^{\infty}P(X=i)$.
כעת, נשים לב שלכל i, $P(X=i)$ מופיע סה"כ בכל הסכימות i פעמים כי i גדול יותר מהמספרים: $k=0,1,2,\dots,i-1$ ולכן עבור כל אחד מהם, $P(X=i)$ יופיע. לכן נשנה את סדר הסכימה באופן הבא:
$\Sigma_{k=0}^{\infty}\Sigma_{i=k+1}^{\infty}P(X=i)=1\cdot P(X=1)+2\cdot P(X=2)+\dots=\Sigma_{i=0}^{\infty}i\cdot P(X=i)$
מכיוון ש X מקבל רק ערכים שלמים אי שליליים אז לפי הנוסחא לחישוב ישיר של התוחלת של X נקבל ש: $E(X)=\Sigma_{i=0}^{\infty}i\cdot P(X=i)$ השווה בדיוק לביטוי שקיבלנו לעיל.

ב. בשאלות מהסגנון הזה נעדיף להציג את הנתונים באמצעות משתנים מקריים, אם יש משהו הקשור לספירה. או באמצעות מאורעות אם מדובר במשהו שקרה או לא קרה. בתרגיל הזה יש לחשב את ההסתברות לקבלת 2 צבעים שונים לכל היותר ולכן מדובר במאורעות.
נגדיר את המאורעות:
R_k - יצא צבע אדום ב k ההטלות הראשונות.
B_k - יצא צבע כחול ב k ההטלות הראשונות.
Y_k - יצא צבע צהוב ב k ההטלות הראשונות.
כעת, אנו רוצים לחשב את ההסתברות שיצא ב k ההטלות הראשונות רק אדום או רק כחול או רק צהוב או אדום וכחול או אדום וצהוב או כחול וצהוב. נחשב תחילה את המקרים שיש בדיוק צבע אחד (נדרוש שיצא הצבע הזה וגם שלא יצאו 2 הצבעים האחרים):
$P(R_k\cap\overline{B_k}\cap\overline{Y_k})=P(\overline{R_k}\cap B_k\cap\overline{Y_k})=P(\overline{R_k}\cap\overline{B_k}\cap Y_k)=\left(\frac{1}{3}\right)^k$ כי אם דורשים את אותו הצבע בכל k ההטלות אזי יש 2 מתוך 6 אפשרויות (2 הפאות הצבועות בצבע המתאים) לקבל את הצבע הדרוש.
עבור 2 צבעים, נחשב (נדרוש שיצאו 2 הצבעים וגם שלא יצא הצבע השלישי):
$P(R_k\cap\overline{B_k}\cap\overline{Y_k})=P(\overline{R_k}\cap B_k\cap\overline{Y_k})=P(\overline{R_k}\cap\overline{B_k}\cap Y_k)=\left(\frac{1}{3}\right)^k$ כי אם דורשים את אותו הצבע בכל k ההטלות אזי יש 2 מתוך 6 אפשרויות (2 הפאות הצבועות בצבע המתאים) לקבל את הצבע הדרוש.
עבור 2 צבעים, נחשב (נדרוש שיצאו 2 הצבעים וגם שלא יצא הצבע השלישי):
$P(R_k\cap B_k\cap\overline{Y_k})=P(R_k\cap\overline{B_k}\cap Y_k)=P(\overline{R_k}\cap B_k\cap Y_k)=\left(\frac{2}{3}\right)^k-2\left(\frac{1}{3}\right)^k$ כי אם דורשים רק את 2 הצבעים בכל k ההטלות אזי יש 4 מתוך 6 אפשרויות (2 הפאות הצבועות בצבע האחד + 2 הפאות הצבועות בצבע השני) לקבל את הצבע הדרוש אבל נוריד את ההסתברות שיצא הכל בצבע אחד מתוך ה 2 (כי כבר חישבנו זאת) ולכן נוריד $\left(\frac{1}{3}\right)^k$ (ההסתברות לצבע אחד) פעמיים (עבור כל צבע).
סה"כ: $P(k)=3\cdot\left(\frac{1}{3}\right)^k+3\left(\left(\frac{2}{3}\right)^k-2\left(\frac{1}{3}\right)^k\right)=3\left(\left(\frac{2}{3}\right)^k-\left(\frac{1}{3}\right)^k\right)$ כאשר $k>0$ (הכפל ב 3 היא כי יש 3 הסתברויות שונות לכל מקרה (לפי הצבע שבחרנו)). אם $k=0$ ההסתברות $P(0)$ שווה ל 1.

ג. מספר ההטלות הנדרשות בניסוי המתואר לעיל הוא מספר אי שלילי ולכן לפי סעיף א', מתקיים:
$E(X)=\Sigma_{k=0}^{\infty}P(X>k)$. כעת, ההסתברות שיהיו יותר מ k הטלות היא שבמהלך ה k ההטלות הראשונות, לא התקבלו כל הצבעים – וזו בדיוק ההסתברות שחישבנו בסעיף ב' (לקבל לכל היותר 2 צבעים). לכן נציב ונקבל: $E(X)=\Sigma_{k=0}^{\infty}P(X>k)=1+\Sigma_{k=1}^{\infty}3\left(\left(\frac{2}{3}\right)^k-\left(\frac{1}{3}\right)^k\right)$. (ההפרדה של 1 מהסכום עבור $k=0$ כי ההסתברות בסעיף ב' עבור $k=0$ הייתה לא לפי הביטוי שמצאנו אלא פשוט 1. מכאן נוציא את 3 מהסכום, נפצל את הסכום ל 2 סכומים ונחשב כל סכום לפי סכום סדרות הנדסיות:
$1+\Sigma_{k=0}^{\infty}3\left(\left(\frac{2}{3}\right)^k-\left(\frac{1}{3}\right)^k\right)=1+3\left(\Sigma_{k=1}^{\infty}\left(\frac{2}{3}\right)^k-\Sigma_{k=1}^{\infty}\left(\frac{1}{3}\right)^k\right)=1+3\left(\frac{\frac{2}{3}}{1-\frac{2}{3}}-\frac{\frac{1}{3}}{1-\frac{1}{3}}\right)=$,
$=1+3\left(2-\frac{1}{2}\right)=1+\frac{9}{2}=\frac{11}{2}$,
משי"ל.