נגזרת	אינטגרל	טור
$(x^n)' = nx^{n-1}$	$\int x^n dx = \frac{x^{n+1}}{n+1}$	$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$
$(\log x)' = \frac{1}{x}, (e^x)' = e^x$	$\int \frac{1}{ax+b} dx = \frac{\log\lvert ax+b\rvert}{a}$	$\sum_{k=0}^{n-1} ar^k = a\frac{1-r^n}{1-r}$
$(\sin x)' = \cos x, (\cos x)' = -\sin x$	$\int xe^x dx = e^x(x-1)$ $\int x^2e^x dx = e^x(x^2-2x+2)$	$\lim_{k\to\infty}\left(1+\frac{a}{k}\right)^k = \sum_{k=0}^{\infty}\frac{a^k}{k!} = e^a$
$\left(f^{-1}(x)\right)' = \frac{1}{f'\left(f^{-1}(x)\right)}$	$\int \log(x)dx = x\log(x) - x$	$\sum_{k=1}^{\infty} ar^k = \frac{ar}{1-r}$
$\left(f(x)g(x)\right)' = f'(x)g(x) + f(x)g'(x)$	$\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx$	
$\left(f\left(g(x)\right)\right)' = f'\left(g(x)\right)g'(x)$		

בחירה עם סדר ללא החזרה	בחירה עם סדר ועם החזרה	בחירה ללא סדר וללא החזרה	בחירה ללא סדר ועם החזרה
$\frac{n!}{(n-k)!}$	n^k	$\binom{n}{k}$	$\binom{n+k-1}{k}$

שם ההתפלגות	PDF	CDF	MGF	תוחלת	שונות
אחידה – בדיד $X\sim U(a,b)$	$\frac{1}{b-a+1}$	$\frac{\lfloor x\rfloor - a + 1}{b-a+1}$	$\frac{e^{as}-e^{(b+1)s}}{(b-a+1)(1-e^s)}$	$\frac{b+a}{2}$	$\frac{(b-a+1)^2-1}{12}$
ברנולי $X\sim Ber(p)$	$p^x(1-p)^{1-x}$		$1-p+pe^s$	p	$p(1-p)$
בינומית $X\sim Bin(n,p)$	$\binom{n}{x}p^x(1-p)^{n-x}$		$(1-p+pe^s)^n$	np	$np(1-p)$
גיאומטרית $X\sim Geo(p)$	$(1-p)^{x-1}p$	$1-(1-p)^k$	$\frac{pe^s}{1-(1-p)e^s}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
היפר-גיאומטרית $X\sim HG(k,n,m)$	$\frac{\binom{m}{x}\binom{n}{k-x}}{\binom{n+m}{k}}$			$\frac{mn}{k}$	
פואסון $X\sim Pois(\lambda)$	$e^{-\lambda}\frac{\lambda^x}{x!}$		$e^{\lambda(e^s-1)}$	λ	λ
אחידה – רציף $X\sim U(a,b)$	$\frac{1}{b-a}$	$\frac{x-a}{b-a}$	$\frac{e^{bs}-e^{as}}{(b-a)s}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$
אקספוננציאלית $X\sim \exp(\lambda)$	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	$\frac{\lambda}{\lambda-s}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
נורמלית $X\sim N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$		$e^{\frac{s^2\sigma^2}{2}+s\mu}$	μ	σ^2

סוג המדגם	אומד נראות מירבית	רווח סמך	בדיקת השערות
$X_1,\dots,X_n\sim Ber(\theta)$	$\hat{\theta}(X) = \frac{1}{n}\sum_{i=1}^{n}X_i = \overline{X}$	בעזרת קירוב נורמלי וחסם על השונות $p \in \left[X \pm \frac{z_{1-\frac{\alpha}{2}}\mathbf{\frac{1}{4}}}{\sqrt{n}}\right]$	$H_0: \theta = \theta_0$ $H_1: \theta = \theta_1 > \theta_0$ $\sum_{i=1}^{n} X_i \geq c$
$X_1,\dots,X_n\sim U(0,\theta)$	$\hat{\theta}(X) = \max\{X_i\}$	$\theta \in \left[max\{X_i\}, \frac{max\{X_i\}}{\sqrt[n]{\alpha}}\right]$	
$X_1,\dots,X_n\sim \exp(\theta)$	$\hat{\theta}(X) = {}^{1}/_{\overline{X}}$		$H_0: \theta = \theta_0$ $H_1: \theta = \theta_1 > \theta_0$

			$\sum_{i=1}^{n} X_i \le c$
$X_1,\dots,X_n\sim N(\mu,\sigma^2)$ כאשר השונות ידועה	$\hat{\mu}(X)=\overline{X}$	$\mu\in\left[\overline{X}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\overline{X}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$	$H_0:\mu=\mu_0$ $H_1:\mu=\mu_1>\mu_0$ $\frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\ge c$ *יש גם מבחן הפוך* *יש גם מבחן דו-צדדי*
$X_1,\dots,X_n\sim N(\mu,\sigma^2)$ *כאשר התוחלת ידועה*	$\widehat{\sigma^2}(X)=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)$		
$X_1,\dots,X_n\sim N(\mu,\sigma^2)$	*שני הקודמים יחד*		
$X\sim N(\mu,\sigma^2)$ דגימה בודדת			$H_0:\mu=\mu_0$ $H_1:\mu>\mu_0$ $\frac{X-\mu_0}{\sigma}\ge c$ כאן β תלוי ב-μ
$X_1,\dots,X_i\sim Pois(\theta)$	$\hat{\theta}(X)=\overline{X}$		$H_0:\theta=\theta_0$ $H_1:\theta=\theta_1>\theta_0$ $\sum_{i=1}^{n} X_i \ge c$

טענה: $P(\bigcup_{i=1}^{n} A_i) = P(A_1) + P(\overline{A}_1 \cap A_2) + \dots + P(\overline{A}_1 \cap \dots \cap \overline{A_{n-1}} \cap A_n)$

<u>הסתברות מותנית</u>: $P(B|A) = \frac{P(A\cap B)}{P(A)}$.

טענה (<u>חוק הכפל</u> או <u>חוק השרשרת</u>): $P(\bigcap_{i=1}^{n} A_i) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1\cap\dots\cap A_{n-1})$.

טענה (<u>נוסחת ההסתברות השלמה</u>): עבור זוג מאורעות A,B מתקיים $P(B) = P(A)P(B|A) + P(\overline{A})P(B|\overline{A})$.

טענה (<u>חוק בייס</u> – Bayes): $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$.

בניסוח אחר: $P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|\overline{A})P(\overline{A})}$.

$A_1,\dots,A_n$ *חלוקה*: $P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$.

הגדרה: אם $P(A|B) = P(B)$ אז נאמר ש-A <u>בלתי-תלוי</u> ב-B. בניסוח אחר, $P(A\cap B) = P(A)P(B)$. נסמן $A \perp B$.

טענה: $A \perp B \Rightarrow A \perp \overline{B}$

הגדרה: נאמר שהמאורעות $A_1,\dots,A_n$ בלתי-תלויים אם לכל $S \subseteq \{1,\dots,n\}$ מתקיים $P(\bigcap_{i\in S} A_i) = \prod_{i\in S} P(A_i)$.

- אבי מטיל n מטבעות ובתיה מטילה $n+1$ מטבעות (הוגנים, ב"ת). מהי הסיכוי שלבתיה יהיה יותר "עץ" מאשר לאבי?
  - נסמן B – לבתיה יש יותר "עץ" מלאבי. נסתכל על n ההטלות הראשונות שלהם ונסמן: A_1 – לבתיה יותר "עץ", A_2 – לאבי יותר "עץ", A_3 – לשניהם אותו מספר הטלות "עץ". נחשב את ההסתברויות של המאורעות:

$P(A_1) = P(A_2),\ \ P(A_3) = 1 - P(A_1) - P(A_2) = 1 - 2P(A_1)$

$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)$

$= 1\cdot P(A_1) + 0\cdot P(A_2) + \frac{1}{2}\cdot\left(1-2P(A_1)\right) = \frac{1}{2}$

הגדרה: <u>פונקציית התפלגות</u> (Probability Mass Function או Probability Distribution) של מ"מ X זו הפונקציה $P_X(x) \coloneqq P(X=x) = \sum_{\omega\in\Omega, X(\omega)=x} P(\omega)$.

הגדרה: <u>ההסתברות לשרוד</u> (Survival Function) על מ"מ $X\sim Geo(p)$ זוהי ההסתברות $P(X>s) = (1-p)^{s+1}$.

טענה (<u>תכונת חוסר הזיכרון</u> של מ"מ גיאומטרי): יהי $Y\sim Geo(p)$, אזי $P(Y>s+t|Y\ge t) = P(Y>s)$.

טענה: יהי $X\sim Bin(p,n)$. הגבול $\lim_{n\to\infty} P(X=x)$ זהה להסתברות $P(Y=x)$ עבור $Y\sim Pois(np)$.

הגדרה: יהי X מ"מ. <u>התוחלת</u> של X היא: $E[X] = \sum_x P(X=x)x$.

טענה: יהי X מ"מ ונגדיר $Y = g(X)$ כאשר $Y:\mathbb{R}\to\mathbb{R}$ פונקציה כלשהי. מתקיים: $E[Y] = \sum_x g(x)P(X=x)$.

- מבצעים n הטלות ב"ת של מטבע עם הסתברות p. מהי ההסתברות שמספר ה"עץ" יהיה זוגי?
  - נגדיר מ"מ $X\sim Bin(n,p)$ הסופר את מספר ה"עצים". נסמן P_n ההסתברות למס' זוגי של "עץ" לאחר n הטלות. $P_0 = 1$. נוסחת הנסיגה היא: $P_n = P_{n-1}(1-p) + (1-P_{n-1})p = p + (1-2p)P_{n-1}$. באינדוקציה, $P_n = \frac{1}{2}(1+(1-2p)^n)$.
  - דרך אחרת – נפתח ישירות:

$$P_n = \frac{1}{2}\left[\sum_{(i=0)}^{n}\binom{n}{i}p^i(1-p)^{n-i} + \sum_{i=0}^{n}(-1)^i\binom{n}{i}p^i(1-p)^{n-i}\right]$$
$$=\frac{1}{2}\left[1+\sum_{i=0}^{n}\binom{n}{i}(-p)^i(1-p)^{n-i}\right]=$$
$$=\frac{1}{2}\left[1+\left(-p+(1-p)\right)^n\right]=\frac{1}{2}(1+(1-2p)^n)$$

טענה: יהי X מ"מ אי-שלילי, אזי $\sum_{x=0}^{\infty}P(X>x)=E[X]$.

טענה (<u>ליניאריות התוחלת</u>): X מ"מ, $E[aX+b]=aE[X]+b$.

הגדרה: ה<u>שונות</u> של מ"מ X $Var(X) \coloneqq E[(X-E[X])^2]$.
<u>סטיית התקן</u> של X היא $Std(X)=\sqrt{Var(X)}$.

הגדרה: <u>המומנט ה-k</u> של מ"מ X הוא $E[X^k]$.

טענה: $Var(X)=E[X^2]-E[X]^2$.

טענה: X מ"מ, $Var(aX+b)=a^2Var(X)$.

הגדרה: X,Y מ"מ. <u>ההתפלגות המשותפת</u>: $P(X=x,Y=y) \coloneqq P(\{X=x\}\cap\{Y=y\})$. <u>ההתפלגות השולית</u> של X: $P_X(x) \coloneqq P(X=x)=\sum_y P_{X,Y}(x,y)$.

טענה (<u>ליניאריות התוחלת</u>): יהיו X,Y משתנים מקריים, אזי $E[X+Y]=E[X]+E[Y]$.

מטילים קוביה עם k פאות, כשלפאה ה-i הסתברות p_i. יהי X_i מספר הפעמים שיצאה הפאה ה-i ב-N הטלות. ה-$k-1$ ערכים הראשונים קובעים את הערך האחרון. זו <u>התפלגות מולטינומית</u>:

$$P_{X_1,\dots,X_k}(n_1,\dots,n_k)=\binom{N}{n_1,n_2,\dots,n_k}p_1^{n_1}p_2^{n_2}\cdots p_k^{n_k}$$
$$=\frac{N!}{n_1!\,n_2!\cdots n_k!}p_1^{n_1}p_2^{n_2}\cdots p_k^{n_k}$$

ההתפלגות השולית של משתנה מסוים:

$$P_{X_i}(n_i)=\sum_{n_1,\dots,n_{i-1},n_{i+1},\dots,n_k}P_{X_1,\dots,X_k}(n_1,\dots,n_k)$$

על כל משתנה אפשר לחשוב כהצלחה/כישלון בבחירת i. $X_i\sim Bin(N,p_i)$.

הגדרה: יהי X מ"מ ו-A מאורע, נגדיר $P_{X|A}(x)=P(X=x|A) \coloneqq \frac{P(\{X=x\}\cap A)}{P(A)}$ <u>ההתפלגות המותנית</u> של X בהינתן המאורע A.

הגדרה: יהיו X,Y מ"מ, נגדיר
$P_{X|Y}(x,y)=P(X=x|Y=y)=\frac{P(\{X=x\}\cap\{Y=y\})}{P(Y=y)}=\frac{P_{X,Y}(x,y)}{P(Y=y)}$ <u>ההתפלגות המותנית</u> של X בהינתן Y.

בניסוח אחר. $P(X=x,Y=y)=P(Y=y)P(X=x|Y=y)$, שזהו חוק הכפל (או חוק השרשרת) עבור משתנים מקריים. באופן דומה אפשר לקבל גם את נוסחת ההסתברות השלמה:

$$P(X=x)=\sum_y P(X=x,Y=y)$$
$$=\sum_y P(Y=y)P(X=x|Y=y)$$

- משדר שולח 1 בהסתברות p ו-0 בהסתברות $1-p$, באופן ב"ת בשידורים קודמים. נגדיר Z – מספר השידורים ליחידת זמן, וידוע ש- $Z\sim Pois(\lambda)$. מה ההתפלגות של ה-1-ים ביחידת זמן?
  - נסמן ב-X את מספר ה-1-ים, Y מספר ה-0-ים ונחשב את ההתפלגות המשותפת:

$$P(X=n,Y=m)=P(X=n,Y=m,Z=n+m)$$
$$=P(X=n,Y=m|Z=n+m)P(Z=n+m)=$$
$$=e^{-\lambda}\frac{\lambda^{n+m}}{(n+m)!}\binom{n+m}{n}p^n(1-p)^m$$
$$P(X=n)=\sum_{m=0}^{\infty}P(X=n,Y=m)$$
$$=\sum_{m=0}^{\infty}e^{-\lambda}\frac{\lambda^{n+m}}{(n+m)!}\binom{n+m}{n}p^n(1-p)^m$$
$$=\sum_{m=0}^{\infty}e^{-\lambda}\frac{\lambda^{n+m}}{n!\,m!}p^n(1-p)^m=$$
$$=\frac{(\lambda p)^n e^{-\lambda p}}{n!}\sum_{m=0}^{\infty}\frac{(1-p)^m e^{-\lambda(1-p)}\lambda^m}{m!}$$
$$=\frac{(\lambda p)^n e^{-\lambda p}}{n!}e^{-\lambda(1-p)}\sum_{m=0}^{\infty}\frac{\left(\lambda(1-p)\right)^m}{m!}=\frac{(\lambda p)^n e^{-\lambda p}}{n!}$$

  - כאשר כאן המעבר האחרון היה לפי טור טיילור של e^x (במקרה הזה, $x=\lambda(1-p)$). לבסוף קיבלנו ש-$X\sim Pois(\lambda p)$, מה שדי סביר אינטואיטיבית.

הגדרה: יהיו X,Y מ"מ ו-A מאורע. <u>התוחלת המותנית</u> של X בהינתן Y או A מוגדרת להיות:

$$E[X|Y=y]=\sum_x xP(X=x|Y=y)$$
$$E[X|A]=\sum_x xP(X=x|A)$$

טענה (<u>נוסחת התוחלת השלמה</u> עבור מאורעות): תהי $A_1,\dots,A_n$ חלוקה של Ω. אזי $E_i\left[E_X[X|A_i]\right]=E[X]$.

טענה (<u>נוסחת התוחלת השלמה</u> עבור משתנים מקריים): יהיו X,Y מ"מ. אזי $E_Y\left[E_X[X|Y]\right]=E[X]$.

טענה: $E[XY]^2\le E[X^2]E[Y^2]$.

טענה: לכל $g(x,y)$, $E[g(X,Y)|Y]=\sum_x g(X,Y)P_{X|Y}(x)$.

טענה: לכל X,Y,Z מתקיים $E[E[X|Y,Z]|Z]=E[X|Z]$.

הגדרה: מ"מ X,Y ייקראו <u>בלתי-תלויים</u> אם לכל x,y מתקיים $P(X=x,Y=y)=P(X=x)P(Y=y)$.

הגדרה: יהיו X,Y,Z מ"מ. המ"מ X,Y ייקראו <u>בלתי-תלויים בהינתן Z</u> אם לכל x,y,z מתקיים $P(X=x,Y=y|Z=z)=P(X=x|Z=z)P(Y=y|Z=z)$.

- יהי Y מ"מ שהוא ההסתברות שהשמש זורחת מחר (אחיד, עם תומך $\left\{0,\frac{1}{n},\dots,\frac{n-1}{n},1\right\}$) ו-$X$ מ"מ הסופר את מספר הזריחות ב-100 הימים הקרובים אז:
  - $X\sim Bin(100,Y)$, ומתקיים $E[X|Y]=100Y$ (תוחלת של מ"מ בינומי).
  - כעת אפשר להשתמש בנוסחת התוחלת השלמה ולקבל:

$$E[X]=E\left[E[X|Y]\right]=\sum_{i=0}^{n}P\left(Y=\frac{i}{n}\right)E\left[X\middle|Y=\frac{i}{n}\right]=\frac{1}{n+1}\sum_{i=0}^{n}\frac{100i}{n}$$

- מטילים מטבע עם הסתברות p לעץ עד לקבלת עץ בפעם הראשונה. לאחר מכן מטילים שוב את אותו מספר הטלות וסופרים כמה עץ יצא. מה ההסתברות שלא יצא עץ בכלל בסיבוב השני?
  - נסמן X – מספר ההטלות בסיבוב הראשון, $X\sim Geo(p)$. נסמן גם Y – מספר ההטלות בסיבוב השני, וכאן יש תלות בין המשתנים: $Y|X=x\sim Bin(x,p)$, כלומר $Y\sim Bin(X,p)$.
  - כעת אנו מחשבים את $P(Y=0)$:

$$P(Y=0)=\sum_{x=1}^{\infty}P(Y=0,X=x)=\sum_{x=1}^{\infty}P(Y=0|X=x)P(X=x)$$

$$=\sum_{x=1}^{\infty}(1-p)^{x}p(1-p)^{x-1}=\frac{1-p}{2-p}$$

- יהיו $X_1\sim Geo(p_1),X_2\sim Geo(p_2)$. אזי $\min\{X_1,X_2\}\sim Geo\big(1-(1-p_1)(1-p_2)\big)$.

הגדרה: יהיו $X_1,\dots,X_n$ מ"מ. הם ייקראו <u>בלתי-תלויים</u> אם $P_{X_1,\dots,X_n}(x_1,\dots,x_n)=\prod_{i=1}^{n}P_{X_i}(x_i)$.

טענה: אי-תלות בשלשות גוררת אי-תלות בזוגות.

- נתון $n\geq 3$ מספר טבעי. בשלב 1 מגרילים K מתוך $\{1,2,\dots,n\}$ בהסתברות לאו דווקא אחידה, אבל יודעים את $\mu=E[K]$. בשלב 2, דוגמים מתוך $\{1,2,\dots,n\}$ K איברים ללא החזרה בהסתברות אחידה, ונסמנם S. נסמן גם X – מספר האיברים ב-S שהם ב-$\{1,2,3\}$.
  - נמצא את התוחלת של X: בהינתן K, נסמן X_i (עבור $i\in\{1,\dots,K\}$) מ"מ המקבל 1 אם המספר ה-i שנבחר הוא ב-$\{1,2,3\}$ ו-0 אחרת. מתקיים $X|K=\sum_{i=1}^{K}X_i|K$, וכן $E[X_i|K]=\frac{3}{n}$ (משתנה בינומי עם פרמטר $\frac{3}{n}$).

$$E[X]=E\left[E[X|K]\right]=E\left[E\left[\sum_{i=1}^{K}X_i|K\right]\right]=E\left[KE[X_1]\right]=E\left[K\frac{3}{n}\right]=\frac{3\mu}{n}$$

  - נמצא את ההתפלגות של X. יש שני סוגים של איברים (1,2,3 וכל השאר), מוציאים K איברים ושואלים מה ההסתברות ש-X מתוכם הם מהסוג הראשון. כלומר, $X\sim HG(K,3,n-3)$.
- מטילים מטבע שהסתברותו היא p, עד שמתקבל רצף של $r>0$ עצים.
  - נסמן X את מספר ההטלות, Y מספר ההטלות עד שהתקבל פלי (לראשונה). כמובן, $Y\sim Geo(1-p)$.

$$E[X]=E\left[E[X|Y]\right]=\sum_{y=1}^{\infty}P(Y=y)E[X|Y=y]=(*)$$

  - על התוחלת המותנית אנו יכולים לומר את הדבר הבא:

$$E[X|Y=y]=\begin{cases}r & y>r\\ y+E[X] & y\leq r\end{cases}$$

  - זאת משום שאם $y\leq r$, אז y ניסיונות כבר קרו, ועכשיו שוב ממשיכים עם X כאילו שחזרנו להתחלה. לכן:

$$(*)=\sum_{y=1}^{r}(y+E[X])(1-p)p^{y-1}+\sum_{y=r+1}^{\infty}r(1-p)p^{y-1}$$

$$=\frac{1-p^{r}}{1-p}+(1-p^{r})E[X]\Rightarrow E[X]=\frac{1-p^{r}}{p^{r}(1-p)}$$

  - כך נוכל לחשב גם את השונות, שכן $E[X^2|Y=y]=\begin{cases}r^2 & y>r\\ E[(y+X)^2] & y\leq r\end{cases}$.

טענה: X,Y מ"מ ב"ת. אזי $E[XY]=E[X]E[Y]$.

טענה: (טריק שימושי) $E[XY]=E\left[E[XY|Y]\right]=E\left[YE[X|Y]\right]$.

טענה: X,Y מ"מ ב"ת. אזי $Var(X+Y)=Var(X)+Var(Y)$.

- n סניפים רוצים להתאחד. בכל סניף i מס' הלקוחות היומי X_i מתפלג פואסון עם פרמטר λ_i באופן ב"ת. מספר הלקוחות בסניף המאוחד, שנסמנו X:
  - התוחלת: $E[X]=\sum_{i=1}^{n}E[X_i]=\sum_{i=1}^{n}\lambda_i$.
  - נפתח את פונקציית ההתפלגות לסניף המאוחד – לצורך הנוחות נסתכל כרגע על $n=2$:

$$\begin{aligned}P(X=x)&=P(X_1+X_2=x)\\&=P((\{X_1=0\}\cap\{X_2=x\})\uplus\dots\\&\quad\uplus(\{X_1=x\}\cap\{X_2=0\})=\\&=\sum_{m=0}^{x}P(X_1=m)P(X_2=x-m)=\\&=\sum_{m=0}^{x}e^{-\lambda_1}\frac{\lambda_1^m}{m!}e^{-\lambda_2}\frac{\lambda_2^{x-m}}{(x-m)!}=\\&=\frac{1}{x!}e^{-(\lambda_1+\lambda_2)}\sum_{m=0}^{x}\frac{x!}{m!\,(x-m)!}\lambda_1^m\lambda_2^{x-m}=\\&=\frac{1}{x!}e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^x\end{aligned}$$

  - כלומר, קיבלנו ש-$X\sim Pois(\lambda_1+\lambda_2)$.
  - כעת, בהינתן $X=s$, מהי התפלגות X_i? בה"כ $i=1$ ואז:

$$\begin{aligned}P(X_1=x|X=s)&=\frac{P(X_1=x,X=s)}{P(X=s)}=\frac{P(X_1=x,X_2+\dots X_n=s-x)}{P(X=s)}\\&=\frac{P(X_1=x)P(X_2+\dots+X_n=s-x)}{P(X=s)}=\\&=\frac{e^{-\lambda_1}\frac{\lambda_1^x}{x!}\,\frac{e^{-(\lambda_2+\dots+\lambda_n)}(\lambda_2+\dots+\lambda_n)^{s-x}}{(s-x)!}}{\frac{e^{-(\lambda_1+\dots+\lambda_n)}(\lambda_1+\dots+\lambda_n)^{s}}{s!}}\\&=\frac{\frac{s!}{x!\,(s-x)!}\lambda_1^x(\lambda_2+\dots+\lambda_n)^{s-x}}{(\lambda_1+\dots+\lambda_n)^x(\lambda_1+\dots+\lambda_n)^{s-x}}=\\&=\binom{s}{x}\left(\frac{\lambda_1}{\sum_{i=1}^{n}\lambda_i}\right)^x\left(\frac{\sum_{i=2}^{n}\lambda_i}{\sum_{i=1}^{n}\lambda_i}\right)^{s-x}\end{aligned}$$

  - וכעת אם נסמן $p=\frac{\lambda_1}{\sum_{i=1}^{n}\lambda_i}$, אז מקבלים ש- $X_1|X=s\sim Bin(s,p)$.

הגדרה: <u>מ"מ רציף</u> הוא פונקציה $X:\Omega\to\mathbb{R}$, כך שקיימת פונקציה $f_X(x)$ המקיימת $P(X\in B)=\int_B f_X(x)dx$, הנקראת <u>פונקציית צפיפות הסתברות</u>, והמקיימת את האקסיומות הבאות:

(1) $f_X(x)\geq 0$ לכל x

(2) $\int_{-\infty}^{\infty}f_X(x)dx=1$

הגדרה: אם X מ"מ רציף ו-f_X פונקציית הצפיפות שלו, ה<u>תוחלת</u> שלו היא: $E[X]\coloneqq\int_{-\infty}^{\infty}xf_X(x)dx$.

לכל פונקציה של מ"מ $g:\mathbb{R}\to\mathbb{R}$, $E[g(X)]=\int_{-\infty}^{\infty}g(x)f_X(x)dx$.

הגדרה: אם X מ"מ רציף ו-f_X פונקציית הצפיפות שלו, ה<u>שונות</u> שלו היא: $Var(X)\coloneqq\int_{-\infty}^{\infty}(x-E[X])^2f_X(x)dx$.

טענה (<u>תכונת חוסר הזיכרון</u>): יהי $X\sim\exp(\lambda)$, אזי $P(X>t+s|X>s)=P(X>t)$.

הגדרה: יהי X מ"מ עם פ' צפיפות $f_X(x)$, לפונקציה F_X הבאה נקרא <u>פונקציית ההתפלגות המצטברת</u> של X: $F_X(x)=P(X\leq x)=\int_{-\infty}^{x}f_X(t)dt$. *מתקיים*: $f_X(x)=\frac{dF_X(x)}{dx}$. *כמו כן*.

(1) $P(a \le X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx$
(2) $F_X(x)$ *לא יורדת*
(3) *מתקיים* $\lim_{x\to-\infty} F_X(x) = 0$
(4) *מתקיים* $\lim_{x\to\infty} F_X(x) = 1$
(5) $F_X(x)$ *רציפה מימין:* $\lim_{\varepsilon\to0^+} F_X(x+\varepsilon) = F_X(x)$ לכל x

- לכל $\{X_i\}_{i=1}^n$ קבוצת מ"מ ב"ת ו-$X = \max\{X_1, \dots, X_n\}$ מתקיים: $F_X(x) = F_{X_1}(x) \cdots F_{X_n}(x)$.
- $X_1, \dots, X_n$ מ"מ רציפים ב"ת ש"ה עם פ' צפיפות $f_X(x)$.
  - אם $Y = \max\{X_1, \dots, X_n\}$ אז $F_Y(y) = \left(F_X(y)\right)^n$.
  - אם $Y = \min\{X_1, \dots, X_n\}$ אז $F_Y(y) = 1 - \left(1 - F_X(y)\right)^n$.
- יהיו $X_i \sim \exp(\lambda_i)$ מ"מ ב"ת, אזי $\min\{X_1, \dots, X_n\} \sim \exp(\sum_{i=1}^n \lambda_i)$.

טענה: יהי X מ"מ רציף אי-שלילי, אזי $E[X] = \int_0^\infty P(X > x)dx$.

הגדרה: $Z \sim N(0,1)$ נקרא נורמלי סטנדרטי, וה-CDF שלו מסומנת $\Phi(z)$. עבור $0 \le \alpha \le 1$ נסמן z_α הערך עבורו $P(Z \le z_\alpha) = \alpha$.

טענה: אם $Z \sim N(0,1)$ אז $aZ + b \sim N(b, a^2)$.

הגדרה: מ"מ X, Y נקראים רציפים במשותף אם לכל $D \subseteq \mathbb{R}^2$ מתקיים: $P\left((x,y) \in D\right) = \iint_D f_{X,Y}(u,v)dudv$, כאשר $f_{X,Y}$ נקראת פונקציית הצפיפות המשותפת שלהם.

פונקציית הצפיפות המשותפת מקיימת את התכונות הבאות:

(1) *לכל* u, v *מתקיים* $f_{X,Y}(u,v) \ge 0$
(2) *מתקיים* $\int_{-\infty}^\infty \int_{-\infty}^\infty f_{X,Y}(u,v)du\,dv = 1$

הגדרה: פונקציית הצפיפות השולית של X אם X, Y רציפים במשותף היא $f_X(u) = \int_{-\infty}^\infty f_{X,Y}(u,v)dv$.

הגדרה: (X,Y) מ"מ. הפונקציה $F_{X,Y}(u,v) = P(X \le u, Y \le v)$ נקראת פונקציית הצטברות משותפת.

כאשר (X,Y) *רציפים במשותף, מתקיים* $F_{X,Y}(u,v) = \int_{-\infty}^u \int_{-\infty}^v f_{X,Y}(s,t)ds\,dt$. *כאשר המ"מ בדידים, מתקיים* $F_{X,Y}(u,v) = \sum_{x\le u}\sum_{y \le v} P_{X,Y}(x,y)$. *עובד גם על מעורבים.*

הגדרה: יהיו X, Y מ"מ רציפים במשותף ו-$g: \mathbb{R}^2 \to \mathbb{R}$, אז התוחלת המשותפת על $g(X,Y)$ מוגדרת להיות $E[g(X,Y)] = \iint_{\mathbb{R}^2} g(u,v)f_{X,Y}(u,v)dudv$.

הגדרה: יהי X מ"מ רציף ו-A מאורע. פונקציית התפלגות מצטברת מותנית של X בהינתן המאורע A מוגדרת להיות $F_{X|A}(u) = P(X \le u | A)$, ופונקציית הצפיפות המותנית מוגדרת להיות $f_{X|A}(u) = \frac{d}{du}F_{X|A}(u)$.

כעת מתקיים $P(X \in \Gamma | A) = \int_\Gamma f_{X|A}(u)du$, לכל $\Gamma \subseteq \mathbb{R}$.

טענה (נוסחת ההסתברות השלמה): אם $A_1, \dots, A_n$ חלוקה של Ω, אזי $f_X(x) = \sum_{i=1}^n P(A_i) f_{X|A_i}(x)$.

טענה: אם $A = \{X \in \Gamma\}$ עבור Γ קטע כלשהו, אזי $f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{P(X\in\Gamma)} & x \in \Gamma \\ 0 & x \notin \Gamma \end{cases}$.

הגדרה: יהיו X, Y מ"מ רציפים, ההסתברות של X בהינתן Y מוגדרת להיות: $P(X \in \Gamma | Y = v) = \lim_{\varepsilon\to0^+} \frac{P(X\in\Gamma, |Y-v| \le \varepsilon)}{P(|Y-v|\le\varepsilon)}$.

- שוברים מקל באורך L באופן מקרי ואחיד בנקודה מסוימת, נסמן Y אורך המקל שנותר. כעת שוברים את המקל שנותר שוב באופן מקרי ואחיד, ונסמן X אורך המקל הסופי. רוצים למצוא את הצפיפות המשותפת $f_{X,Y}$.
  - ראשית נשים לב ש-$Y \sim U(0,L)$, כלומר $f_Y(y) = \frac{1}{L}$. לכן $F_Y(y) = \frac{y}{L}$, וגם $E[Y] = \frac{L}{2}$.
  - כמו כן, $X|Y \sim U(0,Y)$. כעת מהגדרת צפיפות מותנית אנו יכולים למצוא את הצפיפות המשותפת:
$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \Rightarrow f_{X,Y}(x,y) = f_{X|Y}(x,y)f_Y(y)$$
$$= \begin{cases} \frac{1}{L}\frac{1}{y} & 0 \le x \le y \le L \\ 0 & otherwise \end{cases}$$
  - כעת אפשר למצוא את ההסתברות השולית של X:
$$f_X(x) = \int_{-\infty}^\infty f_{X,Y}(x,y)dy = \int_x^L \frac{1}{Ly}dy = \frac{1}{L}[\log(y)]_x^L$$
$$= \begin{cases} \frac{1}{L}(\log L - \log x) & 0 \le x \le L \\ 0 & otherwise \end{cases}$$
  - וכעת אפשר לחשב את התוחלת של X, נשתמש לשם כך שתוחלת מותנית:
$$E[X|Y=y] = \int_{-\infty}^\infty x f_{X|Y}(x,y)dx = \frac{y}{2} \Rightarrow E[X|Y] = \frac{Y}{2} \Rightarrow E\left[E[X|Y]\right]$$
$$= E\left[\frac{Y}{2}\right] = \frac{1}{2}E[Y] = \frac{L}{4}$$
  - באופן דומה אפשר לחשב את $E[XY] = E\left[E[XY|Y]\right] = E\left[YE[X|Y]\right] = E\left[\frac{Y^2}{2}\right]$ וגם את השונות המותנית.

טענה (חוק בייס הרציף): יהיו X, Y מ"מ רציפים, אזי $f_{X|Y}(u,v) = \frac{f_{Y|X}(v,u)f_X(u)}{f_Y(v)}$.

בגרסה אחרת, אפשר לכתוב במפורש את ההתפלגות השולית של Y *ולקבל:* $f_{X|Y}(u,v) = \frac{f_{Y|X}(v,u)f_X(u)}{\int_{-\infty}^\infty f_{Y|X}(v,t)f_X(t)dt}$.

- לנורה יש אורך חיים אקספוננציאלי עם פרמטר Λ, כאשר $\Lambda \sim U(1,1.5)$. בוחנים את המנורה ומודדים אורך חיים y; מה ניתן לומר עכשיו על Λ?
  - נתחיל במציאת הצפיפות המותנית:
$$f_{\Lambda|Y}(\lambda,y) = \frac{f_{Y|\Lambda}(y,\lambda)f_\Lambda(\lambda)}{f_Y(y)} = \frac{2\lambda e^{-\lambda y}}{f_Y(y)} = \frac{\lambda}{e^{\lambda y}}C$$
  - כעת אפשר למצוא את המקסימום של הפונקציה כדי לראות מה ה-λ הסביר ביותר בהינתן y שקיבלנו מהמדידה:
$$\frac{d}{d\lambda}f_{\Lambda|Y}(\lambda,y) = c\left[e^{-\lambda y} + \lambda e^{-\lambda y}(-y)\right] = 0 \Leftrightarrow ce^{-\lambda y}[1 + \lambda y] = 0 \Leftrightarrow \lambda = \frac{1}{y}$$

טענה (כלל בייס המעורב): יהיו X מ"מ בדיד, Y מ"מ רציף, אז מתקיים: $P_{X|Y}(x,y) = \frac{f_{Y|X}(y,x)P(X=x)}{f_Y(y)}$.

הערה: אפשר לכתוב את המכנה בעזרת נוסחת הסתברות שלמה: $f_Y(y) = \sum_k f_{Y|X}(y,k)P(X=k)$, *כדי להימנע מהצורך לחשב ישירות את הצפיפות השולית.*

טענה (כלל בייס המעורב, 2): יהיו X מ"מ בדיד, Y מ"מ רציף, אז מתקיים: $f_{Y|X}(y,x)=\frac{P_{X|Y}(x,y)f_Y(y)}{\int_{-\infty}^{\infty}P_{X|Y}(x,t)f_Y(t)dt}$.

במכנה מופיעה, למעשה, ההסתברות השולית $P_X(x)$.

הגדרה: תוחלת מותנית של מ"מ רציפים X,Y מוגדרת ע"י $E[X|Y=y]=\int_{-\infty}^{\infty}xf_{X|Y}(x,y)dx$.

הגדרה: מ"מ רציפים X,Y נקראים בלתי-תלויים אם $f_{X,Y}(x,y)=f_X(x)f_Y(y)$.

מהגדרה זו אפשר לקבל גם הסתברות משותפת:

$$\begin{aligned}P(X\in A,Y\in B)&=\int_A\int_B f_{X,Y}(x,y)dy\,dx\\&=\int_A\int_B f_X(x)f_Y(y)dy\,dx\\&=\int_A f_X(x)\int_B f_Y(y)dy\,dx=\\&=\left(\int_A f_X(x)dx\right)\left(\int_B f_Y(y)dy\right)\\&=P(X\in A)P(X\in B)\end{aligned}$$

ולכן האי-תלות שקולה בהגדרה גם לגבי התפלגות מצטברת משותפת: $F_{X,Y}(x,y)=F_X(x)F_Y(y)$.

טענה: מ"מ X,Y הם בלתי-תלויים אמם לכל זוג פונקציות φ,ψ מתקיים $E[\varphi(X)\psi(Y)]=E[\varphi(X)]E[\psi(Y)]$.

טענה: X,Y ב"ת ו-$Z=X+Y$: $f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)dx$.

- בהתקפת קניות שולה מבקרת במספר מקרי N של חנויות, ובחנות ה-i מבזבזת סכום מקרי X_i. נניח ש-N הוא שלם חיובי עם פונקציית התפלגות ידועה ו-X_i כולם בעלי אותה תוחלת μ ושונות σ^2, וכל המ"מ בשאלה ב"ת. מהי תוחלת ושונות עלות מסע הקניות?
  - נסמן ב-T את סך הכסף ששולה הוציאה, $T=\sum_{i=1}^N X_i$. כאן מתקיים:

$$E[T|N=n]=E\left[\sum_{i=1}^N X_i\middle|N=n\right]=\sum_{i=1}^n E[X_i|N=n]=\sum_{i=1}^n E[X_i]=n\mu$$

$$E[T]=E\big[E[T|N]\big]=E[N\mu]=\mu E[N]$$

$$\begin{aligned}E[T^2]&=\sum_{n=1}^{\infty}P(N=n)E[T^2|N=n]\\&=\sum_{n=1}^{\infty}P(N=n)E[(X_1+\dots+X_n)^2|N=n]=\\&=\sum_{i=1}^{\infty}P(N=n)E[nX_1^2+n(n-1)X_1X_2]\end{aligned}$$

  - המעבר הזה נכון כיוון שהמ"מ כולם אותו דבר, ויש לנו רק שני סוגים שונים של גורמים – ריבועים של אותו מ"מ, ומכפלות של שני מ"מ שונים. בשני המקרים תוחלות שונות כי X_i,X_j כאן ב"ת אמם $i\neq j$. אפשר להמשיך מכאן:

$$\begin{aligned}&=\sum_{n=1}^{\infty}P(N=n)\Big[E[nX_1^2|N=n]+E[n(n-1)X_1X_2|N=n]\Big]\\&=\sum_{i=1}^{\infty}P(N=n)(n\sigma^2+n(n-1)\mu^2)\end{aligned}$$

טענה (נוסחת השונות השלמה): יהיו X,Y מ"מ. אזי $Var(Y)=E[Var(Y|X)]+Var(E[Y|X])$.

יהיו X מ"מ, $Y=g(X)$, ונתונה $f_X(x)$. רוצים למצוא את $f_Y(y)$ וכל מה שנובע מזה. כדי לעשות זאת, נעבור דרך ההתפלגות המצטברת: $F_Y(y)=P(Y\le y)=P(g(X)\le y)=\int_{x:g(x)\le y}f_X(x)dx$. במקרים מסוימים, זה יהיה מספיק פשוט כדי שנוכל לחשבו, לגזור את התוצאה ולהגיע ל-$f_Y(y)$.

- יהיו $X\sim U(-1,1)$ ו-$Y=\sqrt{|X|}$. הטרנספורמציה הזאת אינה מונוטונית ולכן נצטרך למצוא את מבוקשנו ידנית:

$$\begin{aligned}f_Y(y)=\frac{d}{dy}F_Y(y)=\frac{d}{dy}P(Y\le y)&=\frac{d}{dy}P\left(\sqrt{|X|}\le y\right)\\&=\frac{d}{dy}P(|X|\le y^2)=\frac{d}{dy}P(-y^2\le X\le y^2)\\&=\frac{d}{dy}\frac{2y^2}{2}=2y\end{aligned}$$

טענה: יהי X מ"מ, $Y=g(X)$ כאשר g מונוטונית ממש. אזי $f_Y(y)=f_X\big(g^{-1}(y)\big)\left|\frac{d}{dy}g^{-1}(y)\right|$.

- 2 חברים קובעים להיפגש. כל אחד מאחר באופן ב"ת בזמן אקספוננציאלי עם פרמטר λ. מה הצפיפות של הפרש הזמנים בהם הם מגיעים?
  - נסמן X,Y זמני ההגעה ו-$Z=X-Y$ ההפרש המבוקש. נתחיל מהמקרה שההפרש חיובי:

$$\begin{aligned}F_Z(z)=P(X-Y\le z)&=1-P(X-Y\ge z)\\&=1-\int_0^{\infty}\int_{z+y}^{\infty}f_{X,Y}(x,y)dx\,dy\\&=1-\int_0^{\infty}f_Y(y)\int_{z+y}^{\infty}f_X(x)dx\,dy=\\=1-\int_0^{\infty}\lambda e^{-\lambda y}\int_{z+y}^{\infty}\lambda e^{-\lambda x}dx\,dy&=1-\int_0^{\infty}\lambda e^{-\lambda y}\left[-e^{-\lambda x}\right]_{z+y}^{\infty}dy\\&=1-\int_0^{\infty}\lambda e^{-\lambda y}e^{-\lambda(z+y)}dy=\\=1-e^{-\lambda z}\int_0^{\infty}\lambda e^{-2\lambda y}dy&=1-e^{-\lambda z}\cdot\frac{1}{2}\end{aligned}$$

$$f_Z(z)=\frac{d}{dz}\left\{1-e^{-\lambda z}\cdot\frac{1}{2}\right\}=\frac{\lambda}{2}e^{-\lambda z}$$

  - במקרה שההפרש שלילי, אפשר לראות משיקולי סימטריה שמקבלים:

$$\begin{aligned}F_Z(z)=P(Z\le z)=P(-Z\ge -z)&=P(Z\ge -z)=1-F_Z(-z)\\&=1-\left(1-\frac{e^{\lambda z}}{2}\right)=\frac{e^{\lambda z}}{2}\end{aligned}$$

$$f_Z(z)=\frac{d}{dz}\frac{e^{\lambda z}}{2}=\frac{\lambda}{2}e^{\lambda z}$$

  - זאת משום של-$(X-Y)$ ול-$(Y-X)$ צריכה להיות אותה התפלגות – סימטריה מוחלטת בין המ"מ.

הגדרה: השונות המשותפת (Covariance) של מ"מ X,Y היא: $Cov(X,Y)=E[(X-E[X])(Y-E[Y])]$.

הגדרה: X,Y נקראים בלתי-מתואמים אם $Cov(X,Y)=0$.

נשים לב שמתקיים: $Cov(X,X)=E[(X-E[X])(X-E[X])]=Var(X)$, וכן $Cov(X,Y)=E[XY]-E[X]E[Y]$.

טענה: יהיו X,Y מ"מ ב"ת. אזי $Cov(X,Y)=0$.

ההיפך לא נכון: ייתכנו מ"מ בלתי-מתואמים שאינם בלתי-תלויים.

טענה: $Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)$.

טענה: $Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$.
ובאופן כללי:
$Var(\sum_{i=1}^n X_i)=\sum_{i=1}^n Var(X_i)+\sum_{i\neq j}Cov(X_i,X_j)$.

טענה: $Var(X)Var(Y)\ge Cov^2(X,Y)$.

הגדרה: יהיו X,Y מ"מ. <u>מקדם המתאם</u> שלהם מוגדר להיות $\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$. *מקדם המתאם מקיים:*

(1) $-1\le\rho(X,Y)\le 1$
(2) $X=aY+b \Leftrightarrow |\rho(X,Y)|=1$
(3) חסר יחידות (היחידות של המונה והמכנה מתבטלות)

- יהיו $X\sim U(0,1), Y\sim U(0,1)$ ב"ת. נגדיר $T=Y-X$.
  - נמצא את הצפיפות של T. לשם כך נצטרך לחשב אינטגרל: $P(T\le t)=P(Y-X\le t)=\iint f_{X,Y}(x,y)dydx$ על קבוצת הנקודות שעבורן $y\le t+x$. את זה נצטרך לעשות בשני אזורים שונים:
    - עבור $-1\le t\le 0$:
$$\int_{-t}^0\int_0^{t+x} f_{X,Y}(x,y)dy\,dx=\int_{-t}^0\int_0^{t+x}f_X(x)f_Y(y)dy\,dx=\int_{-t}^0\int_0^{t+x}1dy\,dx$$
$$=\frac{t^2}{2}+t$$
    - עבור $0\le t\le 1$ נעדיף לחשב את השטח של המשולש מעל האזור שמעניין אותנו:
$$1-P(Y-X\ge t)=1-\int_0^{1-t}\int_{t+x}^1 dy\,dx=t-\frac{t^2}{2}+\frac{1}{2}$$
  - את מה שקיבלנו אפשר עכשיו לגזור: $f_T(t)=\begin{cases}1+t & -1\le t\le 0\\ 1-t & 0\le t\le 1\end{cases}$.

הגדרה: יהי X מ"מ. <u>פונקציה יוצרת מומנטים</u> שלו מוגדרת להיות $M(s)=E[E^{sX}]$.

טענה: יהי X מ"מ, $M(s)$ פונקציה יוצרת מומנטים שלו. אזי $E[X^k]=M^{(k)}(0)$.

משפט: המומנטים (מכל סדר) של מ"מ X קובעים את f_X ביחידות.

- אם X,Y מ"מ ב"ת אז $M_{X+Y}(s)=M_X(s)M_Y(s)$.
- אם $Y=aX+b$, אז $M_Y(s)=E\left[e^{s(aX+b)}\right]=e^{sb}E\left[e^{aX}\right]=e^{sb}M_x(sa)$.
- זבובים נכנסים לחדר, ובחדר מלכודת. הזבוב ה-i שנכנס ייתפס במלכודת בהסתברות p, כלומר $X_i\sim Ber(p)$ ב"ת וש"ה. N – מס' הזבובים שנכנסו לחדר, ונתון $N\sim Pois(\lambda)$. Y – מס' הזבובים שנתפסו:
  - ניזכר ש-$M_{X_i}(s)=(1-p+pe^s)$ ו-$M_N(s)=e^{\lambda(e^s-1)}$:
$$M_Y(s)=E[e^{sY}]=E\left[e^{s\sum_{i=1}^N X_i}\right]=E\left[E\left[e^{s\sum_{i=1}^N X_i}\middle|N\right]\right]$$
  - נסתכל על $N=n$ ספציפי:
$$E\left[e^{s\sum_{i=1}^N X_i}\middle|N=n\right]=E\left[e^{s\sum_{i=1}^N X_i}\right]=\prod_{i=1}^n(1-p+pe^s)=(1-p+pe^s)^n$$
  - ונחזור לתוחלת:
$$E[(1-p+pe^s)^N]=\sum_{n=0}^\infty P(N=n)(1-p+pe^s)^n$$
$$=\sum_{n=0}^\infty e^{-\lambda}\frac{\lambda^n}{n!}(1-p+pe^s)^n$$
$$=e^{-\lambda}\sum_{n=0}^\infty\frac{\left(\lambda(1-p+pe^s)\right)^n}{n!}=e^{-\lambda}e^{\lambda(1-p+pe^s)}$$
$$=e^{\lambda p(e^s-1)}$$
  - קיבלנו את ה-MGF של מ"מ פואסון: $Y\sim Pois(\lambda p)$.

טענה: יהיו $X\sim N(\mu_X,\sigma_X^2), Y\sim N(\mu_Y,\sigma_Y^2)$ ב"ת, ו-$Z=X+Y$. אזי $Z\sim N(\mu_X+\mu_Y,\sigma_X^2+\sigma_Y^2)$.

הגדרה: תהי $(X_n)_{n=1}^\infty$ סדרת מ"מ. נאמר שהסדרה (X_n) <u>מתכנסת בהסתברות ל-X</u> אם לכל $\varepsilon>0$, $\lim_{n\to\infty}P(|X_n-X|\ge\varepsilon)=0$.

טענה (<u>אי-שוויון מרקוב</u>): X מ"מ אי-שלילי. אזי $P(X\ge c)\le\frac{E[X]}{c}$.

טענה (<u>אי-שוויון צ'בישב</u>): X מ"מ עם תוחלת μ ושונות σ^2. אזי $P(|X-\mu|\ge c)\le\frac{\sigma^2}{c^2}$.

מסקנה: $P(|X-\mu|\ge k\sigma)\le\frac{\sigma^2}{(k\sigma)^2}=\frac{1}{k^2}$.

משפט (<u>החוק החלש של המספרים הגדולים</u>): תהי (X_n) סדרה של מ"מ ב"ת שווי התפלגות עם תוחלת μ. אזי הסדרה $\left(\frac{X_1+\dots+X_n}{n}\right)$ מתכנסת בהסתברות ל-μ.

הגדרה: סדרה (X_n) של מ"מ <u>מתכנסת בהתפלגות למ"מ X</u> אם $\lim_{n\to\infty}F_{X_n}=F_X$ *(שוויון פונקציות)*.

- יהיו $X_i\sim U(0,1)$ ב"ת ש"ה, ו-$Y_n=\min\{X_1,\dots,X_n\}$.
  - נראה ש-$Y_n\xrightarrow{p}0$ (מתכנסת בהסתברות ל-0).
$$P(|Y_n-0|\ge\varepsilon)=P(\min\{X_1,\dots,X_n\}\ge\varepsilon)=P(X_1\ge\varepsilon,\dots,X_n\ge\varepsilon)$$
$$=\prod_{i=1}^n P(X_i\ge\varepsilon)=(1-\varepsilon)^n\xrightarrow[n\to\infty]{}0$$
  - נגדיר $W_n=nY_n$. סדרה זו לא מתכנסת בהסתברות ל-0:
$$1-F_{W_n}(\varepsilon)=P(W_n\ge\varepsilon)=P(nY_n\ge\varepsilon)=P\left(\min\{X_1,\dots,X_n\}\ge\frac{\varepsilon}{n}\right)$$
$$=\left(1-F_{X_i}\left(\frac{\varepsilon}{n}\right)\right)^n=\left(1-\frac{\varepsilon}{n}\right)^n\xrightarrow[n\to\infty]{}e^{-\varepsilon}$$
  - קיבלנו ש-$F_{W_n}(\varepsilon)\xrightarrow[n\to\infty]{}1-e^{-\varepsilon}$ לכל ε, וזו פונקציית ההתפלגות המצטברת של מ"מ אקספוננציאלי. לכן $W_n\xrightarrow{D}\exp(1)$.

משפט (<u>משפט הגבול המרכזי</u> – Central Limit Theorem): יהיו $X_1,\dots,X_n$ מ"מ בלתי-תלויים שווי התפלגות ונגדיר $S_n=X_1+\dots+X_n, M_n=\frac{S_n}{n}$. נסמן $E[X_i]=\mu, Var(X_i)=\sigma^2$ (סופיים). נגדיר מ"מ חדש $Z_n=\frac{S_n-\mu n}{\sigma\sqrt{n}}$. אזי מתקיים $\lim_{n\to\infty}P(Z_n\le z)=\Phi(z)$, כלומר Z_n מתכנסת בהתפלגות למ"מ נורמלי סטנדרטי, $N(0,1)$.

באמצעות המשפט ניתן לבצע קירוב נורמלי – הממוצע של קבוצת מ"מ ב"ת ש"ה בעלי תוחלת μ ושונות σ^2 ניתן לקירוב ע"י $N\left(\mu,\frac{\sigma^2}{n}\right)$, והסכום ניתן לקירוב ע"י $N(n\mu,\sigma^2 n)$.

- ההסתברות שמערכת תקרוס ביום מסוים היא 5%. בשנה יש 300 ימי עבודה – מהי ההסתברות שהמערכת תקרוס בלא יותר מ-5 ימים בשנה?
  - המ"מ של הקריסה בכל יום ב-$X_i\sim Ber(0.05)$ ומסתכלים על $Y=\sum_{i=1}^{300}X_i$ ורוצים לדעת מה ההסתברות שערכו יהיה קטן או שווה ל-5. נבצע קירוב נורמלי, השונות והתוחלת ידועים:
$$P(Y\le 5)=P\left(\frac{Y-15}{\sqrt{14.25}}\le\frac{5-15}{\sqrt{14.25}}\right)\cong\Phi(-2.65)\cong 0.004$$

הגדרה: אומד (או סטטיסטי) הוא פונקציה $\hat{\theta}: (X_1,\dots,X_n) \to \Theta$, המחזירה עבור מדגם אקראי אומדן – את הערך של הפרמטר θ שמתאים למדגם שהתקבל: $\hat{\theta}(x_1,\dots,x_n) = \theta$.

הגדרה: הטייה של אומד $\hat{\theta}$ מסומנת ומוגדרת $Bias(\hat{\theta}) = E[\hat{\theta}(X)] - \theta^*$, כאשר θ^* הערך ה"אמיתי" של הפרמטר. אומד שהטייה שלו היא 0 נקרא אומד חסר-הטייה.

הגדרה: אומד ייקרא קונסיסטנטי (עקבי) אם $\hat{\theta}(X) \xrightarrow{p} \theta^*$.

תכונות שנרצה שיהיו לאומד:

1. *אם נחזור על התהליך, נקבל תוצאות (אומדנים) דומים – כלומר, $Var(\hat{\theta}(X))$ נמוך*
2. *קונסיסטנטיות*
3. *הטייה קטנה (או לא מוטה)*

הגדרה: פונקציית הפסד (או טעות) של אומד $\hat{\theta}$ מוגדרת $L(\hat{\theta}(X),\theta^*)$ וההפסד (או סיכון) תחת פונקציית הפסד כזאת מוגדר $E[L(\hat{\theta}(X),\theta^*)]$.

הגדרה: Mean Squared Error היא פונקציית הפסד המוגדרת $L(\hat{\theta}(X),\theta^*) = (\hat{\theta}(X)-\theta^*)^2$, וההפסד לפי פונקציה זו מסומן $R := E\left[(\hat{\theta}(X)-\theta^*)^2\right]$.

$$R = Var(\hat{\theta}(X)-\theta^*) + E[\hat{\theta}(X)-\theta^*]^2 = Var(\hat{\theta}(X)) + Bias(\hat{\theta})^2$$

הגדרה: יהיו $X_i \sim f_X(x;\theta)$ (או $P_X(x;\theta)$). אומד הניראות המירבית (Maximum Likelihood Estimator) מוגדר להיות $\hat{\theta}(X) = argmax_\theta f_X(x_1,\dots,x_n;\theta)$.

הגדרה: אומד $\hat{\theta}(X)$ הוא אסימפטוטי נורמלי אם מתקיים $\sqrt{n}(\hat{\theta}(X)-\theta^*) \xrightarrow{D} N(0,\sigma^2_{X_i})$. *כאן $\sigma^2_{X_i}$ השונות של ה"טבע".*

- יהי $\hat{\theta}(X) = \overline{X}$, ונניח $E[X_i]=\theta^*$ ו-$Var(X_i)=\sigma^2_{\theta^*}$ – כלומר אומדים את התוחלת של ה"טבע". כעת לפי משפט הגבול המרכזי, $\frac{n\hat{\theta}(X)-n\theta^*}{\sqrt{n}\sigma_{\theta^*}} \xrightarrow{D} N(0,1) \Longrightarrow \sqrt{n}(\hat{\theta}(X)-\theta^*) \xrightarrow{D} N(0,\sigma^2_{\theta^*})$, ולכן זהו אומד אסימפטוטי נורמלי (לתוחלת).

משפט: MLE הוא קונסיסטנטי ואסימפטוטי נורמלי.

הגדרה: יהי X מדגם מקרי $X_1,\dots,X_n$ הנדגמים באופן ב"ת ש"ה מתוך הטבע, $f_X(x;\theta)$. רווח סמך (Confidence Interval) הוא זוג פונקציות U,V המספקות א"ש $U(X)\le\theta\le V(X)$.

1. *מוצאים פונקציה $g(X;\theta)$ כך שההתפלגות של g אינה תלויה ב-θ (יש לבדוק זאת)*
2. *נחפש a,b כך ש-$P(a\le g(X_i;\theta)\le b) = 1-\alpha$ – זה אמור להיות קל כי g לא תלויה ב-θ*
3. *מחלצים רווח סמך על θ*

- עבור $X_i\sim N(\mu,\sigma^2)$ – כאשר אומדים את μ (σ^2 ידועה):
  - נגדיר $g(X;\theta) = \frac{\overline{X}-\mu}{\sigma}\sqrt{n}$ שהוא מ"מ נורמלי סטנדרטי ואינו תלוי בפרמטר.
  - נסתכל על $P\left(-z_{1-\frac{\alpha}{2}} \le \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \le z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$ – ההסתברות שווה ל-$1-\alpha$ כיוון ש-

$$1-2P\left(Z\ge z_{1-\frac{\alpha}{2}}\right) = 1-2\left(1-P\left(Z\le z_{1-\frac{\alpha}{2}}\right)\right) = 1-2+2\left(1-\frac{\alpha}{2}\right)$$

  - לאחר חילוץ μ, רווח הסמך: $\mu\in\left[\overline{X}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$.

הגדרה: בבדיקת השערות עלינו לבנות מבחן מהצורה $T(X)\ge c$ כאשר X מדגם מקרי ($X_1,\dots,X_n$ ב"ת ש"ה) עבור שתי השערות - H_0 השערת האפס, H_1 השערת האלטרנטיבה, כאשר המבחן דוחה את השערת האפס.

הגדרה: נסמן ב-R את קבוצת המדגמים שעבורם המבחן דוחה את H_0. אז רמת המובהקות (significance) תסומן $\alpha := P_{H_0}((X_1,\dots,X_n)\in R)$ ועוצמת המבחן (power) תסומן $1-\beta$ עבור $\beta := P_{H_1}\left((X_1,\dots,X_n)\in\overline{R}\right)$.

בהינתן מבחן מסוים ו-n, קביעת α קובעת גם את β. אם רוצים להקטין את שניהם, אין ברירה אלא להגדיל את n.

הגדרה: יהי $x_1,\dots,x_n$ מדגם ספציפי ונסמן ב-$T(x_1,\dots,x_n)$ את ערך הסטטיסטי עבור המדגם (המבחן). ה-p-value של המדגם הוא ההסתברות $P_{H_0}(T(X)\ge T(x_1,\dots,x_n))$. *(כאן X הוא מ"מ של כל המדגם, כלומר $X_1,\dots,X_n$).*

ככל שה-p-value גדול יותר, כך המדגם יותר סביר בהינתן H_0. בעצם, ה-p-value של המדגם $x_1,\dots,x_n$ הוא הערך של α אילו היינו בוחרים לקבוע את הסף של המבחן ב-$c = T(x_1,\dots,x_n)$.

טענה: ה-p-value מתפלג אחיד בין 0 ל-1.

הגדרה: מבחן יחס הניראות לבדיקת השערות הוא מבחן מהצורה $L(X) = \frac{P_{H_1}(X_1,\dots,X_n)}{P_{H_0}(X_1,\dots,X_n)} \ge c$.

- יהיו $X_i\sim Pois(\theta)$ וההשערות: $H_0:\theta=\theta_0, H_1:\theta=\theta_1$ עבור $\theta_0<\theta_1$.
  - נבנה כלל דחיה על פי מבחן יחס הניראות:

$$L(X_1,\dots,X_n) = \frac{P_{H_1}(X_1,\dots,X_n)}{P_{H_0}(X_1,\dots,X_n)} = \frac{\prod_{i=1}^n e^{-\theta_1}\frac{\theta_1^{X_i}}{X_i!}}{\prod_{i=1}^n e^{-\theta_0}\frac{\theta_0^{X_i}}{X_i!}} = \prod_{i=1}^n e^{-(\theta_1-\theta_0)}\left(\frac{\theta_1}{\theta_0}\right)^{X_i}$$
$$= e^{(-\theta_1-\theta_0)n}\left(\frac{\theta_1}{\theta_0}\right)^{\sum_{i=1}^n X_i} \ge c$$

  - זו פונקציה מונוטונית עולה של $\sum_{i=1}^n X_i$ (כי $\theta_0<\theta_1$ ולכן הבסיס גדול מ-1), והמבחן הזה שקול למבחן $\sum_{i=1}^n X_i \ge c'$.

משפט (ניימן-פירסון): יהי $L(X)\ge c$ מבחן יחס הניראות עבור שתי השערות, R – תחום דחיית H_0, α,β כרגיל. יהי $X\in S$ מבחן כלשהו, S – תחום דחיית H_0, α',β' כרגיל. אזי $\alpha'\le\alpha \Longrightarrow \beta\le\beta'$.

כלומר, לא ניתן לשפר את α מבלי לפגוע ב-$1-\beta$.

הגדרה: מבחן יחס ניראות מוכלל הוא מבחן בדיקת השערות מהצורה $H_0:\theta\in\Theta_0, H_1:\theta\in\Theta_1$ שסף הדחייה שלו הוא $L(X) = \frac{\max_{\theta\in\Theta_1}P(X_1,\dots,X_n;\theta)}{\max_{\theta\in\Theta_0}P(X_1,\dots,X_n;\theta)}\ge c$.

לעתים אפשר יהיה להשתמש באומד ניראות מירבית שכבר ידוע לנו כדי למצוא את ערך המקסימום במונה/במכנה.