תוכן העניינים

1 טרמינולוגיה ותזכורת ממבוא לסטטיטיקה ותורת ההסתברות

הגדרה 1.1 מרחב מדגם X
מדובר במשנה מקרי עם התפלגות P קרי $X \sim \mathbb{P}$

הגדרה לפי תורת ההסתברות -
יהי Ω מרחב מדגם בדיד . משתנה מקרי (מ״מ) על מרחב זה הוא
הפונק $X : \Omega \longrightarrow \mathbb{R}$ נאמר ש X מוגדר על כל מרחב ההסתברות (Ω, F, P)שזהו מרחב המדגם
הערות

1. נניח כי $X, Y : \Omega \longrightarrow \mathbb{R}$שני מ״מ שמוגדרים על אותו מרחב הסתברות (Ω, F, P)אז גם $X + Y$ או $X \cdot Y$ או $\alpha \cdot X$ הם מ״מ עם טווח שונה שמוגדרים על אותו המרחב

הנחה נוספת היא שאנו לא יודעים מהי $\mathbb{P}$ ושהיא שייכת למשפחה של התפלגיות(Θ) שתלויות בפרמטר θ
קרי $\mathcal{P} = \{\mathbb{P}_\theta, \theta \in \Theta\}$

הגדרה 2.1 סטטיסטי
הדרך להסיק מהו θ בעזרת סטטסטי שהוא משתנה של הנתונים שיש לנו ממרחב המדגם.
סטטיטי הוא משתנה מקרי שמתואר בעזרת הנתונים שיש לנו בלבד ולא בעזרת פרמטרים אותם אנחנו רוצים לאמוד
הדרך להסיק מהי θ מתחלקת ל3 חלקים

1. אמידה נקודתית
2. רווח סמך
3. בדיקת השערות

הגדרה 3.1 $p.m.f$
$probability\, mass\, functions$ פונקציית מסה הסתברותית
$\mathbb{P}_\theta(X = x) = p_\theta(x)$:

הגדרה 4.1 $c.d.f$ פונק התפלגות מצטברת
$\mathbb{P}_X((-\infty, x]) = \mathbb{P}(X \leq x) = F_x(x)$

טענה 5.1 2 דרכים למצוא את $f_x(x)$

1. יהי X משתנה מקרי בעל צפיפות f_x ותהיי $g:\mathbb{R}\to\mathbb{R}$ פונקציה מונ ממש וגזירה
$Y=g(X)$ הוא משתנה מקרי בעל צפיפות ולכל $y\in\mathbb{R}$ מתקיים כי

$$f_Y(s)=\frac{f_X(s)}{|g'(s)|}=f_X\left(g^{-1}(s)\right)\left(g^{-1}\right)'(s)=\frac{f_x(g^{-1}(s))}{g^{'}(g^{-1}(s))}$$

עמוד 155 בספר של אוהד ראה דוגמא תרגיל 12
תמיד נחשב מי הם $g(x)$ $g^{'}(x)$ ו $g^{-1}(x)$ ואז נציב בנוסחא
דוגמא פרטית $x=c+y$ כאשר $y\sim Exp(1)$ אזי $g(x)=c+x$ מונ עולה וגזירה $g^{'}(x)=1$ ומתקיים $g^{-1}(x)=x-c$
כעת $f_x(s)=\frac{f_y(s-c)}{1}=\begin{cases}e^{-(s-c)} & s\geq 0\\ 0 & s<0\end{cases}$

2. דרך שנייה למצוא זה לחשב מהו $F_X(x)=f_x(X\leq x)$ ואז לעשות $f_X(x)=\frac{d}{dx}F_X(x)$

טענה 6.1 טענה חשובה מהסתברות עבור משתנים מקרים בדידים

יהיו X,Y משתנים מקריים שווי התפלגות לא בהכרח מאותו מרחב הסתברות ותהיי $f:\mathbb{R}\to\mathbb{R}$ אזי $f(X)\overset{d}{=}f(Y)$
עמוד 65 בספר של אוהד תורת ההסתברות

טענה 7.1 טענה ממבוא לסטטיטיקה הרצאה 2
תמיד אם $X\in\mathbb{R}^n$ אפשר להגדיר $E(x^2)=||x||^2$ יש לזה תכונה של סמי נורמה כלומר בלי הדרישה של חיוביות בהחלט !
לכן על טענות שלא מתבססות על זה ניתנות לשימוש

טענה 8.1 אינדיקטור $1_{x\in A}$
איניקטור מוגדר כך $P(X=x)=\begin{cases}P(A) & x=1\\ 1-P(A) & x=0\end{cases}$
הפונקציה $1_{x\in A}$ שווה ל1 אם A מתקיים אחרת 0
טריקים

1. נשים לב כי מתקיים $\prod_{i=1}^{n}1_{\{x-\mu\geq 0\}}=1_{\{min\, x_i\geq\mu\}}$ טענה כמסקנה מתרגיל 3

2. אם $x_1,\dots x_n\overset{i.i.d}{\sim}$ אזי $\sum_{i=1}^{n}1_{\{x_j<l\}}-n=-\sum_{i=1}^{n}1_{\{x_j\geq l\}}$ מועד א 2018 מרועי המלך!

1.0.1 פונקציית נראות

טענה 9.1 <u>פונקציית נראות של $X_1,\dots,X_n$</u>
יהי $X_1,\dots X_n$ מדגם מקרי כל דגימה ב"ת בשנייה הדגימות באו מצפיפות $f_\theta(X_i)$ אזי $L(X,\theta)=\prod_{i=1}^{n}f_\theta\left(X_i\right)$
יהי $X_1,\dots X_n$ מדגם מקרי כל דגימה ב"ת בשנייה הדגימות באו מצפיפות $\mathbb{P}_\theta(X_i)$ אזי $L(X,\theta)=\prod_{i=1}^{n}\mathbb{P}_\theta(X_i)$
יהי $X_1,\dots X_n$ מדגם מקרי כל דגימה ב"ת בשנייה הדגימות באו מצפיפות משולבת גם רציפה וגם בדידה אזי
$L(x;\theta):=\prod_{i=1}^{n}f_\theta\left(X_i\right)+\prod_{i=1}^{n}\mathbb{P}_\theta(X_i)$.
הנחה נוספת היא שכל x_j הוא קבוע ממש ביחס לθ
פונקציית נראות יכולה לקבל ערכים גדולים מ1
<u>פונקציית נראות של דגימה בודדת X</u>
שימו לב שכאשר מבקשים ממני למצוא פונקציית נראות של דגימה בודדת לדוגמא $X\sim U([0,\theta])$
פונקציית הנראות היא ה$p.m.f$ כלומר פונקציית ההתפלגות
אם X הוא משתנה מקרי רציף אקבל כי $L(x;\theta):=f(x|\theta)$ כאשר $f(X,\theta)$ היא פונקציית צפיפות $p.d.f$
אם X הוא משתנה מקרי בדיד אקבל כי $L(x;\theta):=p(x|\theta)$ כאשר $p(X,\theta)$ היא פונקציית הסתברות $p.m.f$
אם X הוא משתנה מקרי כך שפונקציית ההתפלגות שלו משלבת מקרים בדידים ורציפים אזי $L(x;\theta):=f(x;\theta)+p(x;\theta)$
נשים לב שלפי ויקיפדיה $L(X,\theta)=P(X|\theta)$ למרות שθ לא ידועה ואנו מנסים לאמוד אותה אנחנו יודעים מי הם X
זה משחק תפקיד האמידה באייסאנית !!
דוגמא להתפלגות בלי פונקציית נראות
נניח ויש לי מטבע עם צד אחד ראש וצד שני θ
כאשר $\mathbb{P}_\theta=\begin{cases}\theta & x=\theta\\ 1-\theta & x=0\end{cases}$
$X_1,\dots,X_n$ דגימות של הטלות מטבע
אזי לא נוכל לבנות לX פונקציית נראות ! כי כל x_j תלוי בθ
נראות של התפלגות ברנולי $X_1,\dots X_n\overset{i.i.d}{\sim}Ber(\theta)$
$$L(x,\theta)=\theta^{\sum\limits_{i=1}^{n}x_i}\cdot(1-\theta)^{n-\sum\limits_{i=1}^{n}x_i}$$
נראות של התפלגות פואסון $X_1,\dots X_n\overset{i.i.d}{\sim}Po(\lambda)$
$$L(\lambda)=\prod_{i=1}^{n}e^{-\lambda}\cdot\frac{\lambda^{x_i}}{x_i!}=e^{-n\lambda}\cdot\frac{\lambda^{[\sum\limits_{j=1}^{n}X_j]}}{\prod\limits_{j=1}^{n}X_j!}$$
$$l(\lambda)=\log(L(A))=-n\lambda+\sum_{i=1}^{n}x_i\log\lambda-\sum_{i=1}^{n}\log\left(x_i!\right)$$

נראות של בינומית $X_1,\dots X_n \overset{i.i.d}{\sim} bin(n,p)$

$$L(p)=\prod_{i=1}^n \mathbb{P}\left(X_i=x_i\right)=\prod_{i=1}^n \begin{pmatrix} n \\ x_i \end{pmatrix} p^{x_i}(1-p)^{n-x_i}$$

$$\ell(p)=(\sum_{i=1}^m X_i)\log\left(\frac{p}{1-p}\right)+n\cdot m\log(1-p)+\log(C)$$

נראות של נורמלית $X_1,\dots X_n \overset{i.i.d}{\sim} N(\mu,\sigma^2)$

$$L\left(\mu,\sigma^2\right)=\frac{1}{(\sqrt{2\pi\sigma^2})^n}\cdot e^{[-\sum\limits_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}]}$$

$$\ell\left(\mu,\sigma^2\right)=\frac{n}{2}log(2\pi)-\frac{n}{2}log(\sigma^2)-\frac{\sum\limits_{i=1}^n (x_i-\mu)^2}{2\sigma^2}$$

נראות של התפלגות אחידה $X_1,\dots X_n \overset{i.i.d}{\sim} U[a,b]$

$$L(a,b)=\frac{1}{(b-a)^n}\prod_{i=1}^n 1_{\{a\le X_i\le b\}}=\frac{1}{(b-a)^n}1_{\{1\le \min X_I\}}\cdot 1_{\{\max X_i\le b\}}$$

נראות של התפלגות גמא $X_1,\dots X_n \overset{i.i.d}{\sim} \Gamma(\alpha,\beta)$

$$\ell(r,\lambda)=-\lambda\sum_{i=1}^n x_i+(r-1)\sum_{i=1}^n \log\left(x_i\right)+r\cdot n\cdot\log(\lambda)-\log(\Gamma(r))$$

נראות של התפלגות מערכית $X_1,\dots X_n \overset{i.i.d}{\sim} Exp(\theta)$

$$l(X,\theta)=n\cdot ln\theta-\theta\sum_{i=1}^n X_i$$

נראות של התפלגות פרטו $X_1,\dots X_n \overset{i.i.d}{\sim}$

$l(X,\theta)=n\cdot ln\theta+n\theta lnc-(\theta+1)\sum\limits_{j=1}^n lnX_j$ ראה פתרון תרגיל 4 שאלה 3

2.0.1 אומדים לפי שיטת המומנטים מסדר 1

יהיו $X_1,\dots X_n \overset{i.i.d}{\sim}$

טענה 10.1 אומדנים בשיטת המומנטים :

מומנט מסדר k $m_k(X)=\mathbb{E}(X^k)$

$E_{\hat\theta}(X^k)=\frac{1}{n}\sum_{i=1}^n X_i^k$

$E[X_i]=\bar X$

טבלה 1 : שיטת מומנטים

התפלגות	אומדן מומנטי
פואסון$X\sim Po(\lambda)$	$\overline{X}=\lambda$
נורמלית $X\sim N\left(\mu,\sigma^2\right)$	$\bar X=\mu.\sigma^2=\frac{1}{n}\sum_{i=1}^n\left(X_i-\overline{X}\right)^2$
אחידה!$X\sim U(0,b)$	$b=2\overline{X}$
$X\sim Gamma($λ(r,	$\hat\lambda=\frac{\overline{X}}{\hat\sigma^2}$.$\hat r=\frac{\overline{X}^2}{\hat\sigma^2}$
אקספ $X\sim Exp(\lambda)$	$\lambda=\frac{1}{\overline{X}}$
$X\sim Bin(n,p)$	$\hat p=\frac{\overline{X}}{n}$

3.0.1 אומדים לפי שיטת נראות מירבית

יהיו $X_1,\dots X_n \overset{i.i.d}{\sim}$

טבלה 2 : אומדים בשיטת נראות מירבית

התפלגות	אומדן מירבי
$X\sim Po(\lambda)$	$\hat\lambda=\overline{X}$
$X\sim N(\mu,\sigma^2)$	$^2=\frac{1}{n}\sum_{i=1}^n\left(x_i-\overline{X}\right)^2,\hat\mu=\overline{X}$
$X\sim U(a,b)$	$\hat a=\min\limits_{i\in[n]}X_i$ $\hat b=\max\limits_{i\in n}X_i$
$X\sim Exp(\theta)$	$\hat\theta=\frac{n}{\sum_{i=1}^n X_i}=\frac{1}{\overline{X}}$
$X\sim Bin(n,p)$	$\hat p=\overline{X}$
$X\sim ber(p)$	$\hat P=\overline{X}$
$X\sim Geo(\frac{1}{\theta})$	$\hat\theta=\overline{X}$

טענה 11.1

4.0.1 התפלגות רב נורמלית ותכונות שלה

הגדרה 12.1 נאמר שוקטור W הוא מתפלג רק נורמלי אם ניתן להציג אותו כך $W=\mu+AZ$ כאשר $\mu\in\mathbb{R}^m$ ו A היא מטריצה $m\times q$ $Z\sim N(0,I_n)$ הוא וקטור נומרלי סטנדרטי
בהיבט האינטואציטיבי כל קורדינטה היא משתנה מקרי נורמלי חד מימדיי כמו בתורת ההסתברות
-הגדרה שקולה לפי ויקיפדיה
$W\sim N(\mu,\sum)$ מתפלג רב נורמי כך ש$W=(w_1,\dots w_n)$ עם וקטור תוחלות $\mu=E(W)=(E(w_1),E(w_2)\dots E(w_n))$
ומטריצת שוניות משותפות $\sum_{i,j}=E[(X_i-\mu_i)(X_j-\mu_j)]=cov(X_i,X_j)$
תכונות בסיסיות

1. $E(W)=\mu\ cov(W)=AA^T$
2. אם $q=m$ כלומר Aהיא מטריצה הפיכה אזי פונקציית הצפיפות של W היא $f_W(w)=(2\pi)^{\frac{-q}{2}}\cdot det(V)^{\frac{-1}{2}}exp(-(w-\mu)^TV^{-1}(w-\mu)/2)$
3. אם $W^1=\mu_1+A^1Z^1$ וגם $W^2=\mu_2+A^2Z^2$ כך ש$A^1A^{1T}=A^2A^{2T}$וגם $\mu_1=\mu_2$ אזי $W^1\overset{d}{=}W^2$ מדובר במספור למעלה ולא בחזקה !
4. אם $W\sim N(\mu,V)$ הוא מתשנה רק נורמי ו$C\in\mathbb{R}^m$ וקטור קבועים אזי $C^Tw\sim N(C^T\mu,C^TVC)$
5. $W\sim N(\mu,V)\iff$ לכל וקטור קבועים $c\in\mathbb{R}^m$ מתקיים כי $c^TW\sim N(C^T\mu,C^TVc)$
   (א) מסקנה מגניבה כל וקטור שהוא מתפלג נורמי אזי כל תת וקטור שלו התפלג נורמי גם כן ! כי אני יכול לבחור סקלר שיאפס איזו קורדינטה שאני רוצה !
      i. בW כך שc^TW תת וקטור של W התפלג נורמלי
   (ב) טרנספרצמיות לינאריות של משתנים שמתפלגים נורמלי הם עדיין התפלגות נורמלית !
6. אם $W\in\mathbb{R}^m$ כך ש$W\sim N(\mu,V)$ו C היא מטריצה $q\times n$ אזי מתקיים כי $CW\sim N(C\mu,CVC^T)$
7. יהיו n משתנים מקריים מתפלגים נורמי קרי $W_1,\dots W_k\sim N(\mu_k,V^k)$ כלומר לכל משתנה מקרי יש μ ו V לא בהכרח זהים !
   (א) והיו $d_1,\dots d_k$וקטורים קבועים אזי מתקיים כי $\sum_{k=1}^K d_kW_k\sim N(\sum_{k=1}^K d_k\mu_k,\sum_{k=1}^K d_k^2V^k)$
8. נניח כי $W\sim N(\mu,V)$ ויהיו 2 קבוצות אינדקסים L_1,L_2 כך ש$Cov(W_r,W_s)=V_{rs}=0\ for\, r\in L_1 and\, s\in L_2$ לכל
   (א) אזי $\{W_r\, r\in L_1\}$ וגם $\{W_s\, s\in L_2\}$ משתנים מקריים בלתי תלויים כלומר $cov(W_s,W_r)=0\Leftarrow$ **אי תלות רק אם המשתנים מתפלגים רב נורמלי !**
   (ב) נשים לב שלא מדובר במטריצת שוניות 0 זהותית ! נניח יש לי את הוקטור $\begin{pmatrix}x_1\\x_2\\y_1\\y_2\\y_3\end{pmatrix}$ אז $cov(X_i,Y_j)=0$ גורר ש$\vec{X},\vec{Y}$ הם וקטורים בלתי תלויים אבל !
   זה לא גורר שx_1,x_2 הם בלתי תלויים או שy_1,y_2,y_3 הם בלתי תלויים לכן אקבל מטריצה שהיא מכילה בלוקים של0 אבל היא לא בהכרח זהותית 0!
   (ג) מסקנה אם X,Y מתפלגים נורמלי וגם $Cov(X,Y)=0$ אזי X,Y הם בלתי תלויים !

טענה 13.1 טריקים על תוחלת$\mathbb{E}(x)$

1. אם X,Y מ"מ ב"ת אזי $\mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y)$
2. $a\mathbb{E}(a)=a$ קבוע
3. תכונות התוחלת המותנית X,Y,Z מ"מ ו$W=f(Z)$ נזכר בהגדרה של תוחלת מותנית

$$f(a)=\begin{cases}E[Y|X=a] & P(X=a)>0\\ 0 & P(X=a)=0\end{cases}$$

   (א) ליני $\mathbb{E}[X+Y|Z]=\mathbb{E}[X|Z]+\mathbb{E}[Y|Z]$
   (ב) תכונת הרכבה $\mathbb{E}[\mathbb{E}[X|Z]|W]=\mathbb{E}[X|W]$
   (ג) מונו אם $X\overset{a.s}{\geq}Y$ אז $\mathbb{E}[X|Z]\overset{a.s}{\geq}\mathbb{E}[Y|Z]$
   (ד) הוצאת החלק הידוע $\mathbb{E}[XW|Z]\overset{a.s}{=}W\mathbb{E}[X|Z]$
   (ה) השמטת החלק שאינו תלוי אם Zב"ת ב- X אז $\mathbb{E}[X|Y]=\mathbb{E}[X]$ כמעט תמיד
   (ו) טענות מתרגיל 6
      i. $E[X|Y,g(Y)]=E[X|Y]$ לכל g פונקציה
      ii. $E[E[X|Y,Z]|Y]=E[X|Y]$

iii. $Var(X)=E[Var(X|Y)]+Var(E[X|Y])$

4. נוסחאת התוחלת השלמה $\mathbb{E}(X)=\sum_{i=1}^{\infty}\mathbb{E}\left(X|A_i\right)\mathrm{P}\left(A_i\right)$

(א) והמקרה הרציף $f_X(x)=\int_{\Omega}f_{X,Y}(x,y)dy=\int_{\Omega}f_{X|Y}(x,y)f_Y(y)dy=\mathrm{E}\left(f_{X|Y}(x,Y)\right)$

5. תוחלת של פונקציה כמשתנה מקרי $E[Y]=\sum\limits_{x\in\mathbb{R}}f(x)P(X=x)$ עבור $Y=f(x)$ במקרה הרציף יהיה אינטגרל !

6. $E[(\sum\limits_{i=1}^{n}x_i)^2]=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}E[x_i\cdot x_j]$ מסקנה מתרגיל 6 !

טענה 14.1 טריקים על שונות $var(x)$

1. $Var(X)=\mathbb{E}(\,(X-\mathbb{E}(X))^2)=\mathbb{E}(X^2)-\mathbb{E}(X)^2$
2. $Var(X+a)=Var(X)$
3. $Var(aX)=a^2Var(X)$ לכן $(\sigma(aX))=|a|\sigma(X)$
4. אם Y מ"מ בעל שונות סופית **בלתי תלוי ב** X **אז** $Var(X+Y)=Var(X)+Var(Y)$
5. אם a קבוע אזי $var(a)=0$
6. אומד חסר הטיה לשונות $S_n^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)$

(א) $S_n^2=\frac{1}{n}\sum\limits_{i=1}^{n}x_i^2-(\overline{X_n})^2$ לפי הנוסחא $var(X)=E(X^2)-[E(X)]^2$

7. אם X,Y משתנים מקריים בלתי תלויים אזי
$Var(XY)=E(X^2)E(Y^2)-E[X]^2E[Y]^2=E[X]^2Var(Y)+E[Y]^2Var(X)+Var(X)Var(Y)$

(א) אם יש תלות אזי
$Var(XY)=E[X^2Y^2]-(E[XY])^2=cov(X^2,Y^2)+E[X^2]E[Y^2]-(E[XY])^2=cov(X^2,Y^2)+[var(X)+(E[X])^2][Var(Y)+(E[Y])^2]$

טענה 15.1 טריקים של Cov

1. $Cov(X,Y)=\mathbb{E}((x-\mathbb{E}(X)(Y-\mathbb{E}(Y)))=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$
2. אם מ"מ ב"ת אזי $Cov(X,Y)=0$
3. $Cov(X,Y)=Cov(Y,X)$
4. $Cov(X+Z,Y)=Cov(X,Y)+Cov(Z,Y)$
5. $Cov(aX,bZ)=abCov(X,Z)$
6. $Cov(X,X)=Var(X)$
7. $Var(\sum\limits_{n=1}^{N}X_n)=\sum\limits_{n,k\le N}Cov(X_n,X_k)=\sum\limits_{n\le N}Var(X_n)+2\sum\limits_{n<k\le N}Cov(X_n,X_k)$.
8. קורלציה באוכלוסיה $\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\overline{y}\right)\left(x_i-\overline{x}\right)=cov(x,y)$
9. $|Cov(X,Y)|\le\sigma_X\cdot\sigma_Y$ א"ש קושי שוורץ !

5.0.1 תכונות בסיסיות של התפלגויות חשובות

טענה 16.1 **התפגלות בינומית**

1. יהי: $\{X_i\}_{i\in[N]}$ משתני ברנולי p ב״ת אזי $\sum\limits_{i\in[N]}X_i\sim Bin(N,p)$
2. חיבור התפלגויות בינומיות:
3. אם $x\sim Bin(N,p)$ וגם $y\sim Bin(M,p)$ ב"ת ! אזי $(x+y)\sim bin(N+M,p)$

(א)

$$\sum_{n=1}^{M+N}X_i\overset{d}{=}X+Y$$

4. נוסחאת סטרלינג

$$\sqrt{2\pi n}\cdot(\frac{n}{e})^n \le n! \le e\cdot(\frac{n}{e})^n$$

התפלגות פואסון

יהי $X\sim Po(\lambda)$

1. אם $Y\sim Po(\eta)$ ב"ת ב $X\sim Po(\lambda)$ אז $(X+Y)\sim Po(\eta+\lambda)$
2. אם$X\sim Po(\lambda)$ וגם $(Y|X)\sim Bin(X,p)$ אז $Y\sim Po(\lambda p)$

התפלגות נורמלית

1. יהיו לי 2 משתנים נורמלים קרי $Y\sim N(\mu_2,\sigma_2^2)$ $X\sim N(\mu_1,\sigma_1^2)$ אזי מתקיים כי

$$Y|X\sim N(\underbrace{\mu_2+\frac{Cov(Y,X)}{var(X)}(X-\mu_1)}_{E(Y|X)},\underbrace{var(Y)-\frac{Cov^2(Y,X)}{var(X)}}_{Var(Y|X)})$$

מסקנה מתרגיל 6 שנובעת המתפלגות דו נורמלית

2. תרגיל 4 שאלה 8 סעיף ד
3. המשך התכונות בויקיפדיה !

התפלגות אחידה רציפה $U(a,b)$

1. נוחסא למומנט n כלומר $E[X^n]=\frac{b^{n+1}-a^{n+1}}{(n+1)(b-a)}$
2. הזזות לינאריות של $U(a,b)$ עדיין התפלגו לינארית
3. $E[X-E(X)]^n=\frac{(b-a)^n(1+(-1)^n)}{2^{n+1}(n+1)}$

טענה 17.1 וקטור של משתנים מקריים
יש לקרוא את הסיכום במופיע ברישומים

טענה 18.1 חציון
יהיו $X_1,\dots X_{2n+1}$ נאמר שהביטוי הוא חציון אם $P(X_i\le med)=P(X_i\ge med)=\frac{1}{2}$
אם יש לי כמות אי זוגית המרחב המדגם שלי אזי $Med=X_i$ כלשהו ! כלומר באופן פורמלי
$Med=X_1\cdot 1_{\{\le x_1\le\}}+X_2\cdot 1_{\{\le x_2\le\}}+\dots+X_{2n+1}\cdot 1_{\{\le x_n\le\}}$ כאשר $\{\le x_i\le\}$ אומר שהוא בדיוק באמצע המדגם ביחס סדר
אם $X_1,\dots X_{2n}$ נאמר כי $Med=\frac{X_n+X_{n+1}}{2}$
מסקנות מתרגיל 4 שאלה 7

טענה 19.1 אי שיויון קושי שוורץ
לכל ζ,η משתנים מקריים מתקיים כי

$$E(\zeta,\eta)^2\le E^2(\zeta)E^2(\eta)$$

שיעור 11 אצל שירה

טענה 20.1 טריק מגניב הסתברות אם יש לי $X_1,X_2\sim Po(\theta)$ בלתי תלויים אזי

$$E[(\frac{x_1-x_2}{2})^2]=E[(\frac{x_1-\theta+\theta-x_2}{2})^2]=\frac{var(x_1)+var(x_4)}{4}$$

יש כאן שימוש 3 טיעונים
$E(x_1x_2)=E(x_1)E(x_2)$ כי x_1,x_2 ב"ת
$var(x_1)=E(x_1-\theta)^2$
$cov(x_1,x_2)=E[(x_1-\theta)(x_2-\theta)]=0$ כי x_1,x_2 ב"ת
משיעור 8 אצל פבל

טענה 21.1 טורי טיילור חשובים

1. $e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!}$

2. $sinh(x)=\sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!}=\frac{e^x-e^{-x}}{2}$

3. $cosh(x)=\sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!}=\frac{e^x+e^{-x}}{2}$

4. $ln(1+x)=\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\cdot x^n$

טענה 22.1 הגישה הבייסאנית לפי מיכה

הגישה הבייסאנית לניתוח מודלים יוצאת מנק הנחה כי הפרמטר בו אני מעוניינים θ הוא משתנה מקרי ואינו ערך קבוע
מבחינה התיאוריה הבייסאנית הנתונים נוצרים באופן הבא
הפרמטר הלא ידוע θ הוא תוצאה של ערך שהתקבל של משתנה מקרי בעל התפגות π
בהינתן הערך θ שהתקבל בשלב הראשון המדגם של תצפיות ממשתנים מקרים בלתי תלויים קרי $X_1,\dots X_n$ בעל צפיפות $f(x|\theta)$
כלומר סיכום המודל הינו

$$\theta\sim\pi$$

$$X_1,\dots X_n|\theta\sim F_\theta$$

ההתפלגות π נקראת התפלגות האפריורית של הפרמטר והיא מכילה את המידע ואת האמונה שלנו על הערכים של המשתנה המקרי לפני שאנו צופים במגדגם.
הגישה הבייסנאנית אומרת כי הסקה על הפרמטר הלא ידוע צריך להיות בעזרת התפלגות האפוסטריוית שמוגדרת להיות

$$\pi(\theta\mid X_1=x_1,\dots X_n=x_n)=\pi(\theta)\prod_{i=1}^{n}\frac{f(x_i\mid\theta)}{P(X_1=x_1,\dots X_n=x_n)}$$

כאשר $P(X_1=x_1,\dots X_n=x_n)$ היא ההתפלגות המשותפת של התצפיות !
ברישומים של מיכה עמוד 33 בפרק של אמידה
בנוסף חוק בייס אומר את הדבר הבא $\frac{P(A|B)\cdot P(B)}{P(A)}=\frac{P(A\cap B)}{P(B)}=P(B|A)$ מהסתברות מותנית !

6.0.1 משפט הגבול המרכזי החוק החלש והחזק שלהמספרים הגדולים

טענה 23.1 משפט הגבול המרכזי CLT
יהיו $x_1,\dots x_n\overset{i.i.d}{\sim}f_x$ כך ש$E[x_i]=0$ וגם $var(x_i)=1$
אזי מתקיים כי

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}x_i\xrightarrow{d}N(0,1)$$

מקורס תורת ההסתברות 1
הגדרה שקולה תחת אותם הנחות $\sqrt{n}\cdot\bar{X}\xrightarrow{d}N(0,1)$
גרסא משוכללת

1. יהיו לי $x_1,\dots x_n\overset{i.i.d}{\sim}f_x$ עם שונות ותוחלת סופית אזי $\sqrt{n}\cdot(\frac{\bar{X}-\mu}{\sigma_{X_1}})\xrightarrow{d}N(0,1)$
כאשר σ זה השונות של X_1 !! ולא של הממוצע האמפירי !
מתורת ההסתברות 1
למעשה זה הגיוני מהחישוב הזה $\sqrt{n}\cdot(\frac{\bar{X}-\mu}{\sigma_{X_1}})=\sqrt{n}\cdot\frac{1}{n}\sum_{i=1}^{n}\frac{x_i-\mu}{\sigma_{X_1}}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\frac{x_i-\mu}{\sigma_{X_1}})$
$(*)$שונות של ממוצע אמפירי היא $\frac{\sigma^2_{X_1}}{n}$

2. יהיו לי $x_1,\dots x_n\overset{i.i.d}{\sim}f_x$ עם שונות ותוחלת סופית אזי $\sqrt{n}\cdot(\bar{X}-\mu)\xrightarrow{d}N(0,\sigma^2_{X_1})$ משיעור 18

3. **שיטת $\triangle$ משפט הגבול המרכזי נשאר נכון תחת הפעלה של פונקציה גזירה ברציפות** תהיי $g:H\to\mathbb{R}$ גזירה ברציפות אזי

(א)

$$\sqrt{n}(g(\bar{X}_n)-g(\mu))\xrightarrow{d}N(0,var(\theta)\cdot[g^{'}(\theta)]^2)$$

4. תהיי $L(X,\theta)=\prod_{i=1}^{n} f(X_i,\theta)$ פונקצית נראות מירבית
בנוסף $I(\theta)=E_\theta(\frac{\partial log f(X_1,\theta)}{\partial\theta})<\infty$
ו $\hat{\theta}=argmax_\theta L(X,\theta)$
אזי נקבל כי $\sqrt{n}(\hat{\theta}-\theta)\xrightarrow{d} N(0,\frac{1}{I(\theta)})$ לכל $\theta\in\Theta$ משבוע **10** פבל !

טענה 24.1 שבוע 8

טענה 25.1 החוק המספרים הגדולים
החוק החלש של המספרים הגדולים $The Weak LLN$
יהיו $X_1,\dots X_n \overset{i.i.d}{\sim} F_x$ סדרה של משתנים מקרים בלתי תלויים ו שווי התפלגות כך ש$E|X_i|<\infty$
אזי $\bar{X_n}\xrightarrow{P}\mu$
בנוסף בשיעור ראינו כי $\bar{X_n}\xrightarrow{L_1}\mu$
תוך שימוש בלמה הבאה שאם נוסיף לתנאי המשפט את הדרישה כי $E|X_i|^2<\infty$ אזי $\bar{X_n}\xrightarrow{L_2}\mu$
שבוע 8 שיעור 16 אצל שירה
הערה קטנה : מאי שיון הלדר מאינפי 3 נסיק כי אם $\zeta_n\xrightarrow{L_p}\zeta$ אזי בהכרח מתקיים כי $\zeta_n\xrightarrow{L_q}\zeta$ לכל $p<q$ הערה של פבל !
החוק החזק של המספרים הגדולים $The LLN$
יהיו $X_1,\dots X_n \overset{i.i.d}{\sim} F_x$ סדרה של משתנים מקרים בלתי תלויים ו שווי התפלגות כך ש$E|X_i|<\infty$
אזי $\bar{X_n}\xrightarrow{a.s}\mu$ מויקיפדיה

7.0.1 פונק יוצרת מומנטים ופונק אופיינית

הגדרה 26.1 פונק יוצרת מומנטים וטענות בסיסיות
יהי X משתנה מקרי הפונק $M_X(t)=E[e^{tx}]$ עבור $t\in(-\delta,\delta)$ סביב הראשית
עבור המקרה הבדיד $M_X(t)=E[e^{tx}]=\sum_{n\in\mathbb{N}}^{\infty} e^{tx}P(X=i)$
עבור המקרה הרציף $M_X(t)=E[e^{tx}]=\int_{-\infty}^{\infty} e^{tx}f_X(x)$
היא פונק יוצרת מומנטים של X השם שלה כי אם גוזרים אותה k פעמים ומציבים 0 מקבלים את המומנט הk
כאשר $E[X^k]$ זה המומנט הk של המשתנה המקרין נשים לב שאם המומטנט בk קיים אזי בהכרח המומנט ה$k-1$ קיים !)
תכונות בסיסיות

1. אם $S_n=\sum_{i=1}^{n} a_iX_i$ כאשר $X_1,\cdots X_n$ משתנים מקרים בלתי תלויים אזי $M_{S_n}(t)=M_{X_1}(a_1t)\cdot M_{X_2}(a_2t)\cdots M_{X_n}(a_nt)$
2. יהי X משתנה מקרי אזי $M_{\alpha X+\beta}(t)=E[e^{(\alpha X+\beta)t}]=e^{\beta t}M_X(\alpha t)$
3. **א"ש צרנוף** לכל $t>0$ ולכל $a\in\mathbb{R}$ $P(X\geq a)\leq M_X(t)\cdot e^{-ta}$
4. **הלמה של הופדינג** יהא X מ"מ המקיים $|X|\leq 1$ וכן $E(X)=0$ אזי לכל $t\in\mathbb{R}$ מתקיים $\mathbb{E}(e^{tX})\leq exp(\frac{t^2}{2})$
5. **א"ש הופדינג** יהיו $\{X_i\}_{i\in[N]}$ משתנים מקריים בלתי תלויים ובעלי תוחלת אפס המקיימים לכל $|X_i|\leq 1$ לכל $i\in[N]$ אזי לכל $a>0$
   (א) אזי מתקיים כי $\mathbb{P}(\sum_{i=1}^{n} X_i\geq a)\leq exp(-\frac{a^2}{2N})$
6. $M_X^K(0)=E[X^K]$ למצוא את המומנט הk זה לגזור k פעמים את הפונק יוצרת מומנטים ולהציב 0
   (א) המוטיבציה לטענה נובעת מטור טיילור $M_X(t)=E[e^{tx}]=1+tE[X]+\frac{t^2E[X^2]}{2!}+\cdots+\frac{t^nE[X^n]}{n!}$בשילוב הטענה כי $E[0]=0$
7. אם ל2 משתנים מקריים X,Y מתקיים כי $M_X(t)=M_Y(t)$ אזי $F_X(x)=F_Y(y)$
8. אם X הוא וקטור של משתנים מקריים אזי $M_X(t)=E[e^{<t,X>}]$
9. הקשר של פונק יוצרת מומנטים לפונק אופיינית היא כזה $\varphi_X(t)=M_X(it)$ כאשר i מספר מרוכב

הגדרה 27.1 פונק אופיינית ותכונות בסיסיות שלה
לכל משתנה מקרי X נגדיר $\varphi_X(t)=E[e^{itX}]$ עבור $t\in\mathbb{R}$
נשים לב שהביטוי תמיד מוגדר היטב היות ו $|e^{itX}|=|cos(tX)+isin(tX)|<1$
תכונות בסיסיות

1. ל2 משתנים מקריים יש את אותה התפלגות כלומר $X_1\overset{d}{=}X_2 \iff \varphi_{X_1}(t)=\varphi_{X_2}(t)$
2. $|\varphi_X(t)|<1$ וגם $\varphi_X(0)=1$
3. המומנטים הk של משתנה מקריX קיימים אזי הם מחושבים כך $E[X^k]=i^{-k}\cdot\varphi_X^{(k)}(0)$ לגזור k פעמים ולהציב 0 כפול i^{-k}
   (א) אם $\varphi_X(t)$גזירה k פעמים בנק 0 אזי כל המומנטים של Xקיימים עד k $E[X^k]=i^k\varphi_X(0)$ ויקיפדיה

4. אם $X_1,\dots X_n$ משתנים מקריים ב"ת אזי $\varphi_{a_1X_1+\dots a_nX_n}(t)=\varphi_{X_1}(a_1t)\cdot\dots\cdot\varphi_{X_n}(a_nt)$
5. אם $Y=aX+b$ עבור $a,b\in\mathbb{R}$ אזי $\varphi_Y(t)=e^{itb}\varphi_X(at)$
6. ויקיפדיה ושיעור 17 פבל

8.0.1 כל האי שוניום החשובים בהסתברות

1. **אי שיון בול** לכל אוסף של מאורעות מתקיים כי $\mathbb{P}(\bigcup_{k=1}^{\infty}A_k)\le\sum_{k=1}^{\infty}P(A_k)$
2. **אי שיון מרקוב** יהי X משתנה מקרי אי שלילי ולכל $a>0$ מתקיים כי $\mathbb{P}(X\ge a)\le\frac{\mathbb{E}(X)}{a}$
   (א) ניסוח שקול $\mathbb{P}(X\ge\mathbb{E}(X)\cdot a)\le\frac{1}{a}$
   (ב) אי שיון מרקוב ההפוך $\mathbb{P}(X\ge\mathbb{E}(X))\ge 0$
3. **אי שיויון צבישיב** יהי X בעל שונות סופית אזי לכל $a>0$ מתקיים כי $\mathbb{P}(|X-\mathbb{E}(X)|\ge a)\le\frac{Var(X)}{a^2}$
   (א) לכל X משתנה מקרי בעל שונות סופית וסטיית תקן σ_X ולכל $b>0$ מתקיים כי $\mathbb{P}(|X-\mathbb{E}(X)|\ge b\sigma_X)\le\frac{1}{b^2}$
4. **אי שיון ינסין** תהיי f פונקציה קמורה כמו (x^2)אם היא מקיימת $f(\alpha a+\beta b)\le\alpha f(a)+\beta f(b)$ אזי מתקיים כי $f(E(X))\le E(f(X))$
   (א) תהיי f פונקציה קעורה אם היא מקיימת $f(\alpha a+\beta b)\ge\alpha f(a)+\beta f(b)$ אזי מתקיים כי $E(f(X))\le f(E(X))$
5. **הכללה של א"ש מרקוב** לכל X משתנה מקרי בעל תוחלת סופית ומומנט k אזי לכל $a>0$ מתקיים כי $\mathbb{P}(|X-\mathbb{E}(X)|\ge a)\le\frac{\mathbb{E}((|X-\mathbb{E}(X)|)^k)}{a^k}$
6. **הכללה של א"ש צבישב** לכל X מ"מ אי שלילי בעל תוחלת סופית ומומנט מסדר k אזי לכל $a>0$ מתקיים כי $P(|X-\mathbb{E}(X)|\ge a\sqrt[k]{var(x)})\le\frac{1}{a^k}$.
7. יהי X מ"מ בעל מומנט מעריכי אזי לכל $t>0$ עבורו $M_X(t)$מוגדרת לכל $a\in\mathbb{R}$ ולכל $a\in\mathbb{R}$ מתקיים $\mathbb{P}(X\ge a)\le M_X(t)e^{-ta}=\mathbb{E}(e^{tx})e^{-ta}$
8. **הלמה של הופדינג** יהא X מ"מ המקיים $|X|\le 1$ וכן $E(X)=0$ אזי לכל $t\in\mathbb{R}$ מתקיים $\mathbb{E}(e^{tX})\le exp(\frac{t^2}{2})$
9. **א"ש הופדינג** יהיו $\{X_i\}_{i\in[N]}$ משתנים מקריים בלתי תלויים ובעלי תוחלת אפס המקיימים לכל $|X_i|\le 1$ לכל $i\in[N]$ אזי לכל $a>0$
   (א) אזי מתקיים כי $\mathbb{P}(\sum_{i=1}^{n}X_i\ge a)\le exp(-\frac{a^2}{2N})$
10. **אי שיויון קושי שוורץ** לכל ζ,η משתנים מקריים מתקיים כי $[E(\zeta,\eta)^2\le E^2(\zeta)E^2(\eta)$
11. **אי שיון המשולש בהסתברות** יהיו ζ,μ משתנים מקריים ותהיי $a\in\mathbb{R}$ אזי $P(|\zeta+\mu|\ge a)\le P(|\zeta|\ge\frac{a}{2})+p(|\mu|\ge\frac{a}{2})$

2 מודלים סטטיסטים

הגדרה 1.2 מודל סטטיסטי
קיימת רשימה של $X\ data$
מניחים שהנתונים הגיעו מהתפלגות כלשהו $\mathbb{P}_{\theta_0}$-
H משפחה של התפלגיות שמאמינים ש$\mathbb{P}_{\theta_0}\in H$
נרצה לדעת מהי ההתפלגות האמתיתית כאשר $X\sim\mathbb{P}_{\theta_0}$
בעיות בסטיטיקה
שיעור 2

1. אמידה נקודתית
2. רווחי סמך
3. בדיקת השערות

הגדרה 2.2 מודל פרמטרי ואי פרמטרי
מודל הוא נקרא פרמטרי כאשר $\mathbb{P}_{\theta_0}\in H$ וH ניתן לשיכון ב$\mathbb{R}^d$ כאשר $d<\infty$
כאשר d הוא מספא דרגות החופש
מודל נקרא אי פרמטרי אם H ניתן לשיכון ב$\mathbb{R}^d$ כאשר $d=\infty$
למשל כאשר צריך לצייר פונקציה ולהגדיר אותה בכל נק על הישר הממשי $\mathbb{R}$
$X_1\dots X_n\sim f$ וצריך לאמוד את f כולה

הגדרה 3.2 מודל ניתן לזיהוי
נאמר כי המודל $(\mathbb{P}_\theta)_{\theta\in\Theta}$ ניתן לזיהוי אם $\mathbb{P}_{\theta_1}(x)=\mathbb{P}_{\theta_2}(x)\ \forall x\in\Omega \Leftarrow \theta_1=\theta_2$
θ
כלומר לכל התפלגות יש רק θ אחת שמתאימה לה אזי המודל ניתן לזיהוי
באופן שקול אם $\theta_1\neq\theta_2 \Leftarrow$ קיימת $A\in\Omega$ כך ש$\mathbb{P}_{\theta_1}(A)\neq\mathbb{P}_{\theta_2}(A)$
שבוע2
שיעור 3

טענה 4.2 מודל ניתן לזיהוי $\iff$ קיים סטטיטאי T כך שהפונקציה $f:\theta\longrightarrow\mathbb{E}_\theta(T)$ היא חד חד ערכית
שבוע2
הוכחה בשבוע 2
שיעור3
מסקנה מהטענה היא שהמודל ניתן לזיהוי אם קיימת f חד חד ערכית
כאשר $f:\theta\to Var(T)$
הערה פרקטית : נביט בדוגמא $X_1,\dots X_n\sim N(\mu,\sigma^2)$ ובסטטיטי $T(X)=\begin{pmatrix}X_1\\X_1^2\end{pmatrix}$
אקבל כי במקרה הנל f היא חחע נשים לב שאם מספר הנעלמים = מספר המשוואות אזי המטריצה המייצגת של מערכת המשוואות תהיה הפיכה קרי f חחע דוגמא בתרגיל כאן

$$ET(X)=\begin{pmatrix}\mu_1\\\sigma_1^2+\mu_1^2\end{pmatrix}=\begin{pmatrix}\mu_2\\\sigma_2^2+\mu_2^2\end{pmatrix}\Rightarrow\begin{cases}\mu_1=\mu_2\\\sigma_1^2+\mu_1^2=\sigma_2^2+\mu_2^2\end{cases}$$

1.0.2 מספיקות ומשפט הפירוק של פישר נוימן

הגדרה 5.2 מספיקות
מוטיבציה
על מנת לבצע הסקה על משתנים מקריים לא תמיד צריך את כל ה$data$ כדי לבצע הסקה כעת נציג לצמצם את כמות ה$data$
יהי T סטטיטי כלשהו נאמר כי הוא **מספיק**
אם ההתפלגות המותנית של X תחת $T(X)$ הוא קבוע שלא תלוי בθ כלומר $\mathbb{P}_{\theta_0}(X=a\,|\,T(X)=b)$ לא תלוי בθ !
קבועה ביחס לθ
שבוע3
שיעור4

טענה 6.2 שימוש בהגדרת המספיקות
יהי $X\sim\mathbb{P}_{\theta_0\in H}$ משתנה מקרי כלשהו

1. נחשב\נמצא סטטיטי $T(X)$ שהוא **מפסיק**
2. נשכח מX
3. אגדיר משתנה מקרי חדש $X^{'}\sim\mathbb{P}_{\theta_0}(X=s|T(X)=t)$ נשים לב כי $X^{'}$ מתפלג בצורה שלא תלויה בθ
4. מתקיים כי $X\overset{d}{=}X^{'}$ איזה מגניב ! הערה $X\neq X^{'}$ אבל יש להם את אותה התפלגות !
5. כלומר עבור כל הסקה שנרצה לעשות הערך $T(X)$ הוא כל מה שאנחנו צריכים מבלי לדעת מהו θ
6. שבוע 3

טענה 7.2 משפט הפירוק של פישר נוימן
תהיי $\mathbb{P}_{\theta\in H}$ מודל סטטיטי עם פונקציית נראות $L(X,\theta)$ ו $X\sim\mathbb{P}_{\theta\in H}$
הסטטיסטי $T(X)$ הוא מספיק $\iff$ קיימים פונקציות $h(x)$ $g(T(X),\theta)$ כך ש

$$L(X,\theta)=g(T(X),\theta)\cdot h(x)\ \ \forall x\in X\ \ \theta\in H$$

הערה : נשים לב כי $h(x)$ לא תלוי בθ וזה קריטי !
שבוע 3

הגדרה 8.2 גסות
נאמר כי $T_1(X)$ הוא גס יותר מ$T_2(X)$ אם קיימת פונקציה כך ש $T_1(X)=\varphi(T_2(X))$
אינטואציה :גס יותר הכוונה שהוא מכיל פחות $data$ למשל $V=\begin{pmatrix}x_!\\\vdots\\\vdots\\x_n\end{pmatrix}$ אז $x_j=\varphi(v)$ קורד ספציפית הוא אומד גס יותר והוא מכיל פחות מידע !
כלומר נוכל להגיע מ$T_1(X)$ ל$T_2(X)$ על ידי מניפולציות אלגבריות
אבל לא נוכל להגיע מ$T_2(X)$ ל $T_1(X)$
נאמר כי $T_1(X)$ ו$T_2(X)$ הם שקולים אם $T_1(X)$ גס יותר מ$T_2(X)$
וגם $T_2(X)$ גס יותר מ$T_1(X)$
הערה : אם φ היא חד חד ערכית ! אזי $T_1(X)$ ו $T_2(X)$ שקולים
מסקנה חשובה אם קיימת f כך ש$f(T_1(X))=T_2(X)$ ו$T_2(X)$ הוא סטטיטי מספיק אזי $T_1(X)$ מספיק גם כן !
שבוע 3
מסקנה חשובה מתרגיל 4
אומד נראות מירבית הוא האומד גס יותר מכל סטטיסטי מספיק
כי $\hat{\theta}=\underset{\theta}{argmax}L(X,\theta)=\underset{\theta}{argmax}[g(T(X),\theta)\cdot h(X)]$
כאשר T הוא סטטיסטי מספיק כלשהו !

הגדרה 9.2 סטטיטי min
נאמר כי $T(X)$ הוא סטטיטי min מספיק אם $T(X)$ גס יותר מכל סטטיטי מספיק אחר שאפשר למצוא .
כלומר $T(x)$ ניתן להגיע אליו על ידי מניפולציות אלגבריות וכל סטטיסטי
שבוע 3

טענה 10.2 קריטריון לסטטיטי min
נכון לשבוע 3 אין לנו קריטריון/ אלגוריתם למצוא את הסטיטסטי המספיק min
אבל אם יש לי T סטטיטי מספיק כלשהו אני יכול להכיע האם הוא min או לא
אם T הוא סטטיטי ויהיו $x, y \in X$ נביט ב $\frac{L(x,\theta)}{L(y,\theta)}$
אם $\frac{L(x,\theta)}{L(y,\theta)}$ קבועה ביחס לθכלומר כשθ זז הביטוי לא משתנה מתקיים כי $T(y) = T(x)$ לכל $x, y \in X$
אזי T הוא סטטיסטי מספיק min !
באופן שקול
אם קיימים $x, y \in X$ כך ש $T(y) \neq T(x)$ אזי $\frac{L(x,\theta)}{L(y,\theta)}$ לא קבועה ביחס לθ כלומר כשθ זז הביטוי משתנה גם כן !
אזי T הוא סטטיטי מספיק min!
הערה פרקטית :
הטענה שקולה יותר קלה הרבה פעמים
נזכר שאם הפונקציה היא קבועה אזי הנגזרת שלה מתאפסת בכל נקודה וזו טענה שימושית
אם מנסים להראות ישירות מהגדרה כי $T(y) = T(x)$
שבוע 3

טענה 11.2 סטטיסטי הmin המספיקים להתפלגיות שונות

1. אם $Po(\lambda) \overset{i.i.d}{\sim} X_1, \ldots X_n$ אזי $T(X) = \sum_{i=1}^{n} X_i$ הוא סטטיטי מספיק min !
2. אם $Gamma(\alpha, \beta) \overset{i.i.d}{\sim} X_1, \ldots X_n$ אזי $T(X) = (\prod_{j=1}^{n} X_j, \sum_{j=1}^{n} X_j)$ הוא סטטיסטי מספיק min לα, β
3. אם $Ber(\theta) \overset{i.i.d}{\sim} X_1, \ldots X_n$ אזי $\sum_{i=1}^{n} X_i$ הוא סטטיסטי מספיק min!
4. אם $N(\mu, \sigma^2) \sim X_1, \ldots X_n$ אזי $S(X) = (\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2)$ הוא סטטיסטי מספיק min ! לμ ו σ^2

1.2 אלמנטים בתורת ההחלטות

הגדרה 12.2 הsetup הוא כזה יש לי $X_1, \ldots X_n \sim P_{\theta \in H}$ כאשר הטבע יודע מהו θ
והסטטיסקאי לא יודע מהו θ והוא בוחר באומדים לθ קרי $\hat{\theta_1, \ldots}$
הסטטיסטקאי רוצה לבנות פונקציית הפסד על מנת למזער נזקים !
הפסד של סטטיסקאי זה רווח של הטבע ולהפך יש כאן מעיין משחק סכום אפס
לדוגמא ראה שיעור 7 אצל פבל או שבוע 4 דוגמא עם תפוזים !

הגדרה 13.2 פונקציית הפסד רגעי $l : \theta \times \theta \to \mathbb{R}$
כך $l(\theta, \theta(X))$ כלומר ערך אחד הוא θ פרמטר לא ידוע
$\theta(X)$ זה האומד לθ שמתבסס על הנתונים X
פונק הפסד נפוצות
אם $\theta \in \mathbb{R}^n$ וμ הוא אומד לθ

$$l(\theta, \mu) = (||\theta - \mu||_p)^p = [\sqrt[p]{(\theta_i - \mu_i)^p}]^p$$

$$l(\theta, \mu) = Max_i|\theta_i - \mu_i|$$

אותם נורמות שראינו באינפי 3 קרי נורמת p $|| \cdot ||_p$ ונורמת $|| \cdot ||_\infty$
שבוע 4 שיעור 7

הגדרה 14.2 פונקציית סיכון
$R : \Theta \times D \to \mathbb{R}$
D מרחב כל האומדים לθ
כאשר

$$R(\theta, \theta(X)) = E_\theta l(\theta, \theta(X))$$

כלומר ההפסד הממוצע זו התוחלת של פונק הפסד רגעי

האומד עם פונקציית הסיכון הקטנה ביותר הוא אומד מעולה ! כלומר אם $R(\theta,\bar{\theta}(X)) < R(\theta,\tilde{\theta}(X))$
אזי נעדיף את $\bar{\theta}$ על פני $\tilde{\theta}$אציין כי ־ זה סתם סימון ולא מייצג כאן ממוצע !!
שיעור 7 שבוע 4
בהמשך אנחנו רואים כי לא ניתן להשוואת בין פונקציות סיכון באופן כללי
ראה שיעור 8 שבוע 4 דוגמא עם פואסון כשלוקחים 2 אומדים קרי
אם $X_1,\ldots X_n \sim Po(\theta)$ היות ו$Var(\theta) = E(X) = \theta$
נביט ב2 האומדים $\hat{\theta}_1 = \bar{X}_n$ וגם $\hat{\theta}_2 = \frac{1}{n}\sum\limits_{j=1}^{n}(X_j - \bar{X}_j)^2$

הגדרה 15.2 אומד לא כשר
נאמר כי $\tilde{\theta}$ הוא אומד לא כשר אם קיים אומד $\hat{\theta}$ כך ש$R(\hat{\theta}_n,\theta) \le R(\tilde{\theta}_n,\theta)$ לכל $\theta \in H$
וקיים $n \in \mathbb{N}$ כך שאקבל א"ש חריף !
הערות :
פרקטית אם יש לי אומד $\hat{\theta}(X)$ שלא מבוסס על כל התצפיות אני יכול לבנות אומד יותר טוב ממנו כך
אם $R = m.s.e$ אזי אגדיר $\bar{\theta}(X) = \frac{1}{n}\hat{\theta}(X)$
ואקבל כי $R(\bar{\theta},\theta) = \frac{1}{n^2}var(\hat{\theta}) + \frac{1}{n}E[\hat{\theta}] \le var(\hat{\theta}) + E[\hat{\theta}] = R(\hat{\theta},\theta)$
מסקנה מהלמה של ראו בלאק ווי
אם $M(X)$ הוא סטטיסטי מספיק min ביחס ל$R = m.s.e$ **ולא** קיימת פונקציה φ כך ש$\hat{\theta}(x) = \varphi(M(X))$ אזי $\hat{\theta}(X)$ לא כשר !
שבוע 4 שיעור 7

הגדרה 16.2 אומד כשר
הוא אומד כשר אם לא קיים אומד $\hat{\theta}$ כך ש$R(\hat{\theta}_n,\theta) \le R(\tilde{\theta}_n,\theta)$ לכל $\theta \in H$
ולא קיים $n \in \mathbb{N}$ כך שאקבל א"ש חריף !
שבוע 4 שיעור 7
בהמשך רואים כי לא ניתן למצוא את האומד האופטימלי
בשבוע 4 שיעור 8

טענה 17.2 נניח כי $X \sim N(\theta,1)$ ואבחר $\hat{\theta} = c$ אזי $\hat{\theta}$ הוא אומד כשר
מסקנה מהטענה היא שכשרות היא לא תכונה טובה במובן הזה שהיא לא יודעת להבדיל בין אומדים טובים
לבין אומדים לא טובים
שבוע 4 שיעור 7

2.2 3 גישות לבחור באומד האופטימלי

בשיעור 8 אנו רואים 2 דברים חשובים

1. לא ניתן להשוואת בין פונקציות סיכון באופן כללי
2. לא ניתן למצוא את האומד האופטימלי

לכן אנו רוצים לשנות את הצורת חשיבה שלנו עד כה ולהצביע על 3 גישות לבחור אומד אופטימלי

1. גישת $minmax$ מינימקס הסטטיסקאי הפסימי
נגדיר פונק סיכון $r(\hat{\theta}) = \underset{\theta\in H}{Sup}R(\hat{\theta},\theta)$ אנו נעדיף את $\hat{\theta}$ על פני $\tilde{\theta}$ אם $r(\hat{\theta}) < r(\tilde{\theta})$
האומד האופטימי בגישה הזו הינו $\hat{\theta_*}$ כך ש$r(\hat{\theta_*}) \le r(\hat{\theta})$
   (א) אמידה בייסאנית (בפרק של אמידה לקרוא מהי בכללי הגישה בייסאנית בסטטיטיקה)
   פונק הסיכון לפי גישה זו הינה $r_{\pi_1}(\hat{\theta}) = \int\limits_H R(\theta,\hat{\theta})\cdot\pi_1(\theta)d\theta$ המוטיבציה לנוסחא הזו היא נוסחת ההסברות
   כלומר פונק הסיכון כאן מסתכלת על הסיכום המממוצע של θ
   כאשר הסטטסטאי מניח כי θ_0 נבחר מתוך התפלגות ידועה π_1 למעשה θ_0 הוא משתנה מקרי.
   2 הגישות אללו דומות בכך שהם מביאות ערך נומרי מספרי ל$\hat{\theta}$.
   כלומר יש יחס סדר מלא בגישה הזו וכל 2 אומדים הם ברי השוואה כלומר $r(\hat{\theta})$ ו $r(\tilde{\theta})$ הם ברי השוואה
2. להצטמצם לאומדים חסרי הטייה
   (א) מסתבר שאם אני מצמצם את H שלי לאומדים לחסרי הטייה לθ בלבד ניתן באופן מפתיע למצוא את האומד האופטימלי
   החסרון בשיטה הזו שיש הרבה אומדים טובים שאנו מתעלמים מהם למשל אומד בשיטת נראות מירבית הוא לרוב חסר הטייה .
   חוסר הטייה משמעותו שהשגיאה שלי מתפזרת באופן סיטרי באיזשהו מובן.
3. אומדים אוקיוריינטיים

1.2.2 אמידה בייסאנית

הקדמה :
אמידה בייסאנית (בפרק של אמידה לקרוא מהי בכללי הגישה בייסאנית בסטטיטיקה)
פונק הסיכון לפי גישה זו הינה $\theta, r_{\pi_1}(\hat{\theta}) = \int_H R(\theta, \hat{\theta}) \cdot \pi_1(\theta) d\theta$ המוטיבציה לנוסחא הזו היא נוסחת ההסברות
כלומר פונק הסיכון כאן מסתכלת על הסיכום הממוצע של θ
כאשר הסטטסטאי מניח כי θ_0 נבחר מתוך התפלגות ידועה π_1 למעשה θ_0 הוא משתנה מקרי.

טענה 18.2 למצוא מהו $\hat{\theta}$שממזער את הפונקציה $r_{\pi_1}(\hat{\theta})$
אנו רוצים לעבור על פני כל $\hat{\theta} \in H$ ולמזער את הפונקציה $r_{\pi_1}(\hat{\theta}) = \int_H R(\theta, \hat{\theta}) \cdot \pi_1(\theta) d\theta$

ונקבל כי

$$\hat{\theta_\pi}(x) \in argmin_{\mu \in H} \int_H l(\theta, \hat{\theta}) f_{\theta|X}(\theta, x) d\theta$$

שבוע 5 שיעור 9
אמירה כללית כל $r_{\pi_1}(\hat{\theta})$: נשים לב שהסיבה שלא יכולתי לגזור את הביטוי $r(\pi, \hat{\theta})$ נעוץ בעובדה שאני לא יודע איך גוזרים פונקציה

אני יודע איך גוזרים פונקציה שהיא סלקר ! ו$\hat{\theta}$היא לא סלקר !

- **אומד בייס ביחס לפונק הפסד כלשהו הוא תמיד גס יותר מסטטיסטי מספיק** מסקנה מתרגיל 5

טענה 19.2 נשים לב כי $f_{\theta|X}$ הינה

$$f_{\theta|X} = \frac{L(X, \theta)\pi(\theta)}{\int_H L(X, t) \cdot \pi(t) dt}$$

נשים לב שמשוואה זו היא בדיוק חוק בייס $P(B|A) = \frac{P(A \cap B)}{P(B)} = \frac{P(A|B) \cdot P(B)}{P(A)}$ מהסתברות מותנית !
שבוע 5 שיעור 9
נשים לב שמכיון ש$\int_H L(X, t) \cdot \pi(t) dt = f_x(x)$ וכעת $f_x(x)$ לא תלוי בθ היות ועברנו על כל הθ שיש ניתן לכתוב

$$f_{\theta|X} = \frac{L(X, \theta)\pi(\theta)}{\int_H L(X, t) \cdot \pi(t) dt} = C \cdot L(X, \theta)\pi(\theta) =\propto L(X, \theta)\pi(\theta)$$

כאשר C הוא קבוע כלשהו זה נובע מכך ש$f_x(x)$ שהוא המכנה בביטוי לעיל כאשר אני מתנה בx הוא **מספר קבוע וידוע !**
מסקנות מתרגיל 5

טענה 20.2 אומד בייסנאי שממזער לפי שיטת MSE
אם $l(\theta, \mu) = (\theta - \mu)^2$ כלומר $R(\theta, \mu) = E_\theta(\theta - \mu)^2 = Var_\theta(\theta) + (\underbrace{E(\theta) - \mu}_{bias})^2 = MSE$
אזי האומד בייסנאי ש$\hat{\theta}$ממזער את $r_{\pi_1}(\theta)$ הינו $E(\theta|X)$
קרי $\hat{\theta_\pi^*}(X) = E(\theta|X) = \underset{\theta \in H}{argmin}[r_{\pi_1}(\theta)] = \underset{\theta \in H}{argmin}[\int_H (\theta - \mu)^2 \cdot f_{\theta|X}(\theta, X) d\theta]$
אם מקבשים לחשב מהו $E[\theta|X] = \int_{Supp(\theta)} \theta \cdot f_{\theta|X} d\theta = \frac{\int_{Supp(\theta)} \theta \cdot L(\theta, X) \cdot \pi(\theta) d\theta}{\int_{Supp(\theta)} L(X, \theta) \cdot \pi(\theta) d\theta}$
שיעור 9 שבוע 5

טענה 21.2 אומד בייסאני שממזער $E(\theta - \mu|)$

$$\frac{1}{2} = F_{\theta|X}(\mu, X) = \hat{\theta_\pi}(x) = \underset{\theta \in H}{argmin}[\int_H |\theta - \mu| \cdot f_{\theta|X}(\theta, X) d\theta]$$

שיעור 9 שבוע 5

חישוב אמידה בייסאנית על סמך משפחות צמודות

הגדרה 22.2 משפחה של פריורים קרי π נאמר שהיא צמודה למשפחה $L(\theta, X)$
אם $f_{\theta|X} \sim \pi$
כלומר הם מתפלגים באותה התפלגות עם פרמטרים שיכולים להיות שונים !
כאשר כעת π , $f_{\theta|X}$נקראים צמודי התפלגות או צמודי נראות (זה לא קשור לפונקציית נראות)
וπ נקרא צמוד של $f_{\theta|X}$
שיעור 9 שבוע 5 יש להסתכל כל הדוגמא !

טענה 23.2 פרופורציה
כשפותרים פרקטית שאלות הרבה פעמים כדאי לשים לב כי $f_{\theta|X}(x) = C \cdot f_g(x)$ כלומר שההתפלגות הפוסטריות היא קבוע כפול התפלגות מוכרת למשל בינומית פואסונות וכו. נסמן זאת $f_{\theta|X} \propto f_g(x)$
נשים לב שבמקרה מתקיים כי $P(X=\mu) = C \cdot f_g(x) \Rightarrow X \sim f_g(\alpha, \beta)$ כלומר α, β הם פרמטרים כלשהם !
דוגמא אם $P(X=\mu) = C \cdot \underbrace{\binom{n}{x} p^x (1-p)^{n-x}}_{Bin(n,p)}$ אזי Xמתפלג בינומית גם כן עם פרמטרים $Bin(..\cdot, \cdot..)$ כלשהם לא בהכרח שווים ל p, n !
<u>הסיבה לכך נובעת שאני יכול לבצע פעולות אלגבריות כלשהם שיכניסו את הקבוע שלי לביטוי מהצורה של התפלגות בינומית</u>
וההוכחה לך נובעת מתורת החבורות אני יכול להגדיר קבוצה $G = \{Bin(n,p) \mid n, 0 \leq p \leq 1 \in \mathbb{R}\}$ עם כפל רגיל
מדובר כאן בחבורה כשהאיבר ניטרלי הוא 1 בנוסף אני מראה כי $c \in \mathbb{R}$ אזי $c \in G$ כלומר אני יכול לבטא כל סלקר עם G !
והרי זה הגיוני שההתפלגות הבינומית הטווח שלה הוא $\mathbb{R}^+$ ואז מתקיים כי סלקר כפול כל החבורה אז פרמוטציה של החבורה לכן $g \cdot G = G$.
ואז למעשה זה אומר שקיימת סדרה של פעולות אלגבריות שיכניסו לי את הקבוע ! גם אם אני לא יודע להצביע על סדרה כזו ספציפית
כעת אני יכול לומר כי X מתפלג $f_g(x)$ כלשהי עם פרמטרים שאני מצאתי ולומר שהפרמטרים אללו נכונים עד כדי נרמול בקבוע !
כלומר בכתיב רשמי כותבים שמדובר בהתפלגות כלשהי עד כדי נרמול בקבוע
<u>כאשר מוצאים אומד בייס ל$f_{\theta|x}$</u>אנחנו מחפשים את האומד הmin ! שממזער לי את $r_\pi(x)$ כך שגם אם אני יודע מהי $f_{\theta|X|}$ רק עד כדי נרמול האומד הmin שממזער לי את $r_\pi(x)$ נשאר אותו אומד עד כדי נרמול בקבוע של $f_{\theta|X|}$.
כדי למצוא את המנרמל שלי אני יכול פשוט לכתוב $\int_H C \cdot f_{\theta|X} = 1$ ולבודד את C והוא המנרמל שלי.

2.2.2 אומדים חסרי הטייה

הגדרה 24.2 אומד חסר הטייה
נאמר ש$\hat{\theta}$ הוא אומד חסר הטייה לθ אם $b(\hat{\theta}, \theta) = E[\hat{\theta}] - \theta = 0$ עבור $\theta \in H$
הערה לא מחייבת בדרך אומד חסר הטיה הוא ממוצע אמפירי
שבוע 6 שיעור 10

- הרבה פעמים ניתן לקחת אומדים ולהמיר אותם להיות חסרי הטייה

טענה 25.2 לפי שיטת MSE מתקיים כי $E_\theta(\hat{\theta} - \theta)^2 = Var_\theta(\hat{\theta}) + b^2(\hat{\theta}, \theta)$ שבוע 6 שיעור 10

הגדרה 26.2 אומד $UMVUE$ הוא אומד חסר ההטייה הטוב ביותר
נאמר כי $\hat{\theta^*}$ הוא אומד $UMVUE$ אם מתקיים כי $R(\theta, \hat{\theta^*}) = Var_\theta(\hat{\theta^*}) \leq Var_\theta(\hat{\theta}) = R(\theta, \hat{\theta})$ לכל $\theta \in H$ ולכל $\hat{\theta}$אומד כלשהו לθ !
בנוסף אם קיים אומד שהוא $UMVUE$ אזי הוא יחיד עד כדי $a.s$
כלומר אם T_1, T_2 הם אומדי $UMVUE$ אזי מתקיים כי $T_1 \overset{a.s}{=} T_2$
שבוע 6
שיעור 11

-

טענה 27.2 קיימת דוגמא כך שלא קיים לה אומר חסר הטייה
אם $X_1, \ldots X_n \overset{i.i.d}{\sim} ber(\theta)$ נגדיר $q(\theta) = \frac{\theta}{1-\theta}$ אזי לא קיים אומד חסר הטייה ל$q(\theta)$!
מסקנה למעשה התפלגות של X לא משנה כאן **מספיק שהאומד לT יהיה פולינום של θ ממעלה סופית**
כדי לסתור את הטענה שהוא אומר את $q(\theta)$ שהרי $q(\theta) = \frac{\theta}{1-\theta} = \theta \cdot \sum_{i=1}^{\infty} \theta_i$.
נשים לב כי אם $X_1, \ldots X_n \overset{i.i.d}{\sim} Exp(\theta)$ אזי $T(X_1) = e^{-X}$ הוא אומד חסר הטייה ל $q(\theta)$
כי כאן אני מבצע אינטגרציה בחישוב $E[T(X_1)]$!
שבוע 6

טענה 28.2 יהיו $X_1, \ldots X_n \overset{i.i.d}{\sim} f_x$ ויהי $T(X) = \sum_{i=1}^{n} c_i X_i$ עבור $c_i \in \mathbb{R}$
כאשר T הוא סטטיסטי שרוצה לאומד את μ התוחלת של X_i
אם אדרוש כי T יהיה חסר הטייה
אזי בהכרח מתקיים

1. T חסר הטייה $\iff \sum_{i=1}^{n} c_i = 1$
2. $\bar{X} = min\{Var(T) \mid \sum_{i=1}^{n} c_i = 1\}$

כלומר $\bar{X}$ הוא האומד בעל השונות הmin מבין כל האומדים הליניאריים לתוחלת של X_i
מסקנה מתרגיל 6

טענה 29.2 הלמה של $rao-blackwell$
יהי S סטטיטי מספיק ו T אומד ל$q(\theta)$ עבור q פונקציה ידועה
עם $ET^2(X)<\infty$
נגדיר $T^*(X)=E(T(X)|S(X))$
אזי מתקיים כי

$$E[T^*(X)]=T(X)$$

$$Var(T^*(X),q(\theta))\leq Var(T(X),q(\theta))\Rightarrow R(T^*,q(\theta))\leq R(T,q(\theta))$$

בהבט הפרקטי זה אומר שיש לי כאן מתכון איך לשפר אומדים !
יש כאן משחק עם 2 אומדים T,S שבחירה נבונה שלהם יוצרת לי את האומד $UMVUE$
את S אני אעדיף לבחור כסטטיטי מספיק min !
מסקנות מהלמה

1. אומד שאינו גם מסטטיסטי min מספיק הוא לא כשר לפי $m.s.e$ מזה זה אומר:
אם אסמן את $M(x)$ הוא סטטסטי min מספיק **ולא** קיימת פונקציה φ כך ש$T(x)=\varphi(M(X))$ אזי $T(x)$ לא כשר ביחס ל$m.s.e$!
2. ניתן לשפר אומדים חסרי הטייה לפי $R-B$!

חזרה על תוחלת מותנית
נשים לב כי זאת ההגדרה של תוחלת מתונית כמתשנה מקרי

$$f(a)=\begin{cases}E[Y|X=a] & P(X=a)>0\\ 0 & P(X=a)=0\end{cases}$$

כלומר הפלט שלי הוא $E[Y|X=a]$ או 0 והסיכוי לפלט הוא $P(X=a)$.
כמובן שבלמה לעיל אני מתעניין באיך נראה הפלט שלי שהוא בעצמו משתנה מקרי
שיעור 11
שבוע 6

חסם קרמר ראו ושימושיו $Cramer-rao$

טענה 30.2 חסם קרמר ראו $Cramer\ Rao$
מוטיבציה
אם הייתי מטיל מטבע כמות סופית של פעמים אף פעם אני לא אדע במדויק מי זה θ ברמת דויק מושלמת
וכמו שאני לא אדע לעולם מי זה θ ברמת דיוק מושלמת לא אצליח למצוא את האומד הטוב ביותר לפונקציית הסיכון $R(\theta,\hat{\theta})$.
לפי שיטת $m.s.e$ קרי $R(\theta,\hat{\theta})=var(\hat{\theta})+E(\hat{\theta}-\theta)^2$
חסם קרמר- מקבל אומד כלשהו ואומר לי עד כמה האומד שלי יכול להיות נמוך !
למשל אם חסם קרמר מקבל אומד חסר הטייה הוא מחזיר לי פלט מהי פונק הסיכון הנמוכה ביותר שאני יכול לקבל.
ניסוח המשפט
נניח כי המודל הסטטסטאי הוא $(\mathbb{P}_\theta)_{\theta\in H}$ כאשר H הוא מרחב מפרמטים ממימד אחד ! קרי $H\subset\mathbb{R}$!
כאשר $f_{\theta,x}$ היא פונקצית צפיפות גזירה בθ כך ש$x\in\mathbb{R}^n$ ו$\theta\in H$
ו$T(X)$ הוא אומד לθ בנוסף אנו מניחים כי אינטגרציה וגזירה הם ברי החלפה כלומר
$\frac{d}{d\theta}\int\limits_{\mathbb{R}^n}f(x,\theta)dx=\int\limits_{\mathbb{R}^n}\frac{\partial}{\partial\theta}f(x,\theta)dx$
$\frac{\partial}{\partial\theta}\int\limits_{\mathbb{R}^n}T(x)f(x,\theta)dx=\int\limits_{\mathbb{R}^n}T(x)\frac{\partial}{\partial\theta}f(x,\theta)dx$
אזי פונקצית סיכון לפי $m.s.e$ של הסטטיסטי T מקיימת כי

$$R(\theta,\hat{\theta})=Var_\theta(T)+b^2(\theta,T)\geq\frac{(\frac{d}{d\theta}b(\theta,T)+1)^2}{I_{\vec{X}}(\theta)}+b^2(\theta,T)$$

נשים לב שמשתמע מכך ש

$$Var_\theta(T)\geq\frac{(\frac{d}{d\theta}b(\theta,T)+1)^2}{I_{\vec{X}}(\theta)}$$

כאשר $b(\theta,T)=E(T)-\theta$
ו$I_{x_1}(\theta)=E_\theta(\frac{\partial}{\partial\theta}log\cdot f(X,\theta))^2=\int\limits_{\mathbb{R}^n}(\frac{\partial}{\partial\theta}logf(x,\theta))^2f(x,\theta)dx$ ביטוי זה נקרא אינפורמצית פישר !
השיון לעיל לא טריוואלי וההוכחה מופיעה בויקיפדיה
נשים לב כי אם יהיו $f(x,\theta)\overset{i.i.d}{\sim}X_1,\ldots X_n$ אזי $I_{\vec{X}}(\theta)=n\cdot I_{X_1}(\theta)$
הערות

1. אם T הוא אומד חסר הטייה אזי $R(\theta,T)=Var(T)\geq\frac{1}{I(\theta)}$

2. כאשר מדובר במשתנים בדידים אזי נדרוש את אותם הנחות כאשר נחליף את $\int\limits_{\mathbb{R}}$ ב$\sum\limits_{\mathbb{R}}$ ואת פונק הצפיפות בפונק ההסתברות !

3. אם קיבלנו ש T מקיים את כל התנאים הדרושים בשאלה אבל הוא אומד ל$g(\theta)$ למשל $\frac{1}{\theta}$ עבור g פונקציה כלשהי נצטרך בחישוב פונק הסיכון $R(\theta,\hat{\theta}(x))$ להוסיף ולהחסיר את θ באופן מלכותי כדי להשתמש בחסם $Cramer-rao$ ראה תרגול חזרה של רחל תרגיל 7 שאלה 2

טענה 31.2 שבוע 7

הגדרה 32.2 אומד יעיל
נאמר שאומד T הוא יעיל אם הוא מקיים את התכונה ש$R(\theta,T)=cramer-rao-bound$
הערות
נשים לב כי בקרב אומדים חסרי הטייה אומד יעיל הוא אומד $UNVUE$
אבל לא להפך ! אומד $UNVUE$ הוא לא תמיד אומד יעיל !!
ראה דוגמא 2.74 בשבוע 7
כאשר יש אפשרות להראות כי $1_{\{x=0\}}$ הוא אומד $UMUVE$ (בכלים שלא למדנו)
אומדים שמגיעים לחסם $Cramer\ Rao$ כלומר אומדים שהם $UNVUE$

1. אם $Po(\theta)\overset{i.i.d}{\sim}X_1,\dots X_n$ אזי $\overline{X}$ הוא אומד $UNVUE$ ל θ ויקיפדיה

טענה 33.2 טענות בסיסיות על אינפורמציית פישר $I_x(\theta)$ ועל ציון הפילוג
$V=\frac{\partial}{\partial\theta}lnf(X,\theta)$ הנגזרת לפי θ של לוג פונק צפיפות שימוש ישיר של זה
זה נגזרת הנגזרת לפי θ של לוג פונק נראות !

1. $E[V]=E[\frac{\partial}{\partial\theta}lnf(X,\theta)]=0$

2. יהיו x,y משתנים מקרים בלתי תלויים אזי $I_{(x,y)}(\theta)=I_x(\theta)+I_y(\theta)$ כאשר (x,y) הוא וקטור של משתנים ב"ת !
   (א) מסקנה יהיו $f(x,\theta)\overset{i.i.d}{\sim}X_1,\dots X_n$ אזי $I_{\vec{X}}(\theta)=n\cdot I_{X_1}(\theta)$

3. נניח כי $f(x,\theta)$ גזירה פעמים וגם ניתן להחליף את האינטגרציה וגזירה קרי $\frac{\partial^2}{\partial\theta^2}\int\limits_{\mathbb{R}^n}f(x,\theta)dx=\int\limits_{\mathbb{R}^n}\frac{\partial^2}{\partial\theta^2}f(x,\theta)dx$
   אזי מתקיים כי
$$I(\theta)=-E_\theta\frac{\partial}{\partial\theta^2}logf(x,\theta)=-\int\limits_{\mathbb{R}^n}f(x,\theta)\frac{\partial^2}{\partial\theta^2}logf(x,\theta)dx$$

4. $I(\theta)=E_\theta(\frac{\partial}{\partial\theta}log\cdot f(X,\theta))^2=\int\limits_{\mathbb{R}^n}(\frac{\partial}{\partial\theta}logf(x,\theta))^2f(x,\theta)dx$

5. שבוע 7

3.2 התכנסות של משתנים מקריים

הקדמה
בשביל הנושא הזה אנו צריכים להזכר בסדרות של משתנים מקריים מתורת ההסתברות
בנוסף בכל הנושא של הכתנסות במידה שווה של פונקציות מאינפי 2
שבוע 8

הגדרה 34.2 התכנסות נקודתית של משתנים מקריים
תהיי $X_n:\Omega\to\mathbb{R}$ נאמר שהיא מתכנסת נקודתית אם מתקיים $X(\omega)\longrightarrow X_n(\omega)$ לכל $\omega\in\Omega$
שיעור 13 זה חומר מאינפי 2
דוגמא לפונקציה שלא מתכנסת נק
נניח כי $X_n\sim N(0,1)$ ואגדיר $Y_n=\frac{X_n}{n}$ אזי Y_n לא מתכנסת בנקודתית כי אם $X_n=(1,-2,3-,4\cdots)$ אזי $Y_n=1,-1,1-1,\cdots$ וזו לא סדרה מתכנסת !

הגדרה 35.2 התכנסות במידה שווה של משתנים מקריים
תהיי $f_n:\mathbb{R}\to\mathbb{R}$נאמר שהיא מתכנסת במידה שווה אם $\lim\limits_{n\to\infty}\sup\limits_{x\in I}|f_n(x)-f(x)|\longrightarrow0$ אזי $f_n\overset{||\cdot||_\infty}{\longrightarrow}f$
שיעור13
שבוע 8

טענה 36.2 התכנסות במידה שווה גוררת התכנסות נק

הגדרה 37.2 התכנסות בL_p של משתנים מקריים
עבור $P \geq 1$ ζ משתנה מקרי עם $E(\zeta) < \infty$
וגם $E|\zeta_n|^p < \infty$
אם $\lim\limits_{n\to\infty} E|\zeta_n - \zeta|^p = 0$ אזי נאמר כי $\zeta_n \xrightarrow{L_p} \zeta$
שיעור 13 שבוע8
דומגא לפונקציה שלא מתכנסת בL_p אם ζ מתפלגת התפלגות קושי אזי $E(\zeta_n) = \infty$
א"ש שיכולים לעזור להראות התכנסות בL_p :

1. אי שיון צבישב $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{Var(X)}{a^2}$
2. א"ש קושי שוורץ למקרה L^2 : $E(\zeta \cdot \eta)^2 \leq E^2(\zeta)E^2(\eta)$
3. מאי שיון הלדר מאינפי 3 נסיק כי אם $\zeta_n \xrightarrow{L_p} \zeta$ אזי בהכרח מתקיים כי $\zeta_n \xrightarrow{L_q} \zeta$ לכל $p < q$ הערה של פבל !

הגדרה 38.2 התכנסות בהסתברות
תהיי סדרה של מתשנים מקריים ζ_n נאמר שהיא מתכנסת לζ בהתסברות אם
לכל $\varepsilon > 0$
מתקיים כי $\lim\limits_{n\to\infty} [P(\zeta_n \geq \zeta + \varepsilon) + P(\zeta_n \leq \zeta - \varepsilon)] = 0 \Leftrightarrow \lim\limits_{n\to\infty} P(|\zeta_n - \zeta| \geq \varepsilon) = 0$ ונסמן $\zeta_n \xrightarrow{p} \zeta$
התכנסות בL_p גוררת התכנסות בהסתברות !
א"ש שיכולים נפוצים לעזור להראות התכנסות בהסתברות :

1. אם ζ_n משתנה מקרי בדיד ניתן לכתוב כי $P(|\zeta_n - \zeta| > \varepsilon) = P(|\zeta_n - \zeta| > \varepsilon, \zeta_n \in \{value \geq 0\}) \leq P(\zeta_n \in \{value \geq 0\})$
ראה תרגיל 8 שאלה 5
2. אי שיון מרקוב $P(|\zeta_n - \zeta| > \varepsilon) \leq \frac{E(|\zeta_n - \zeta|)}{\varepsilon}$
3. אי שיון צבישב $P(|\zeta_n - E[\zeta_n]| > \varepsilon) \leq \frac{var(\zeta_n)}{\varepsilon^2}$

הגדרה 39.2 בערך כל הא"ש של הסתברות !

הגדרה 40.2 התכנסות בהתפלגות
נאמר כי $X_n \xrightarrow{d} X$ אם מתקיים כי $P(X_n \leq x) = F(X_n) \overset{n\to\infty}{\longrightarrow} F(X) = P(X \leq x)$

הגדרה 41.2 שבוע 8

התנאים הבאים שקולים (משפט הארנק בצרפתית)

1. $\zeta_n \xrightarrow{d} \zeta$
2. $E[g(\zeta_n)] \overset{n\to\infty}{\longrightarrow} E[g(\zeta)]$ לכל g פונקציה רציפה וחסומה
   (א) מסקנה אם $\zeta_n \xrightarrow{d} \zeta$ אזי $g(\zeta_n) \xrightarrow{d} g(\zeta)$ לכל g פונק רציפה !
3. $\varphi_{X_n}(t) = E[e^{it\zeta_n}] \longrightarrow E[e^{it\zeta}] = \varphi_X(t)$
4. שבוע8
5. שיעור 15 שירה

1.3.2 סדר בראש על כל ההתכנסיות השונות היחסים בינהם

1. התכנסות במידה שווה ⇐ התכנסות נק .
   (א) התכנסות במידה שווה ⇐ התכנסות בL_p⇐התכנסות בהסתברות ⇐ התכנסות בהתפלגות .
   (ב)
$$\begin{array}{ccccc} \xrightarrow{L^s} & \underset{s>r\geq 1}{\Rightarrow} & \xrightarrow{L^r} & & \\ & & \Downarrow & & \\ \xrightarrow{a.s} & \Rightarrow & \xrightarrow{p} & \Rightarrow & \xrightarrow{d} \end{array}$$
2. **התכנסות בהסתברות והתכנסות בL_p משמרות אריתמטיקה של גבולות**
   (א) $\zeta_n + \mu_n \xrightarrow{p} \zeta + \mu$ כאשר ζ, μ **הם משתנים מקריים**
   (ב) לכל $c \in \mathbb{R}$ מתקיים כי $c\zeta_n \xrightarrow{p} c\zeta$
   (ג) $\zeta_n \cdot \mu_n \xrightarrow{p} \zeta \cdot \mu$
3. אם $\zeta_n \xrightarrow{p} \zeta$ ותהיי g פונקציה רציפה אזי $g(\zeta_n) \xrightarrow{p} g(\zeta)$ **באופן דומה** אם $\zeta_n \xrightarrow{d} \zeta$ ותהיי g פונקציה רציפה אזי $g(\zeta_n) \xrightarrow{d} g(\zeta)$

4. תהיי $\zeta_n \xrightarrow{d} \zeta$ ותהיי α_n סדרת מספרים כך ש$\alpha_n \longrightarrow \infty$ אזי $\frac{\zeta_n}{\alpha_n} \xrightarrow{p} 0$

5. **למת סלוצקי** תהיי $\zeta_n \xrightarrow{p} c$ כאשר c **קבוע** ו$\mu_n \xrightarrow{d} \mu$ וμ **משתנה מקרי** כלשהו אזי מתקיים כי

(א) שימו לב לנקודה העדינה כאן ברור שיש כאן אריתמטיקה של גבולות עבור התכנסות בהתפלגות אבל כאן יש טענה חזקה יותר שבא $\zeta \xrightarrow{p} \zeta_n$.

(ב) $(\zeta_n, \mu_n) \xrightarrow{d} (c, \mu_n)$

(ג) $\zeta_n + \mu_n \xrightarrow{d} \zeta + \mu$

(ד) $\zeta_n \mu_n \xrightarrow{d} c\mu$

(ה) $\frac{\mu_n}{\zeta_n} \xrightarrow{d} \frac{\mu}{c}\, if\, c \neq 0$

6. נניח כי $\zeta_n \xrightarrow{d} c$ כאשר c **קבוע** ! אזי מתקיים כי $\zeta_n \xrightarrow{p} c$!

7. אם $x_1, \dots x_n$ מתשנים מקרים ב"ת אזי $S_n \xrightarrow{p} S \iff S_n \xrightarrow{a.s} S$ לפי ויקיפדיה

8. אם $X_n \xrightarrow{a.s} X$ וגם $|X_n| < Y$ וגם $E(Y) < \infty$ אזי $X_n \xrightarrow{L^1} X$ לפי ויקיפדיה

9. שבוע 8

10. שיעור 15 שירה

4.2 אמידה במדגמים גדולים

כל הסטטיסטיקה היישומית היא אסימפטוטית כי בשביל סטטיסטיקה טובה צריך הרבה $data$
אנו נדבר על סדרה של תצפיות $X_n = X_1, \dots$ כאשר $X_1, \dots X_n \overset{i.i.d}{\sim} P_{\theta \in H}$
וסדרה של אומדים $\hat{\theta}_n(X_{1:n})$

הגדרה 42.2 סדרה עקיבה
נאמר שסדרה היא עקיבה ל θ אם $\underset{n\to 0}{lim} P(|\hat{\theta_n} - \theta| > \varepsilon) = 0\ \forall \varepsilon > 0 \iff \hat{\theta_n} \xrightarrow{p} \theta \iff \underset{n\to\infty}{lim}[P(\zeta_n \geq \zeta + \varepsilon) + P(\zeta_n \leq \zeta - \varepsilon)] = 0$
באופן דומה נאמר שסדרה עקיבה ל$q(\theta)$ אם $q(\hat{\theta_n}) \xrightarrow{p} q(\theta)$ לכל q פונקציה כלשהי
שיעור 16 שירה סוף שבוע 8
שיעור 18
נזכר בטענה אם כי $\zeta_n \xrightarrow{d} \theta$ כאשר θ קבוע אזי $\zeta_n \xrightarrow{p} \theta$
וכאן בא לידי ביטוי הקשר בין סדרה עקיבה לבין סדרה עקיבה בקצב α_n

הגדרה 43.2 סדרה עקיבה בקצב α_n
תהיי α_n סדרה מונטונית עולה שואפת ל∞ סדרה לא מקרית !
ותהיי $\hat{\theta}_n$ סדרה של אומדים עם התפלגות אסימפטוטית $F(X, \theta)$
נאמר שהסדרה עקיבה בקצב α_n אם $\alpha_n(\hat{\theta}_n - \theta) \xrightarrow{d} \zeta$ כאשר $\zeta \sim F(X, \theta)$
תחת P_θ לכל $\theta \in H$ כלומר $\underset{n\to\infty}{lim} P_\theta(|\alpha_n(\hat{\theta}_n - \theta)| \leq x) = F(x, \theta)$
נאמר שהשגיאה של הסדרה הינה $\hat{\theta} - \theta$ והיא שואפת אסימפוטוטית בקצב $\frac{1}{\alpha_n}$
דוגמא:
תחת תנאי המשפט הגבול המרכזי $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2_{X_1})$
הערה הקצב α_n היא סדרה יחידה !

מוטיבציה להגדרה : ראינו כי לפי החוק החזק/ חלש של המספרים הגדולים מתקיים כי $\bar{X} \overset{a.s/p}{\longrightarrow} E(x_i)$
אבל מסתבר שאם אתה מסתכל על $\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} N(0, \sigma^2_{X_1})$. ומכאן בא המוטיציה להגדיר עקיבה בα_n התכנסות בהתפלגות !
שבוע 10
טענה עקיבות נשמרת תחת הפעלת פונקציה גזירה ברציפות (שיטת $\triangle$)
תהיי $\hat{\theta}_n$סדרת אומדים במבוססת על התצפיות $x_1, \dots x_n$ שאסימפטוטית מתפלגת נורמלית עם שונות של האומד ! $\hat{V}(\theta)$
נשים לב כי $\hat{V}(\theta)$ זה אומד לשונות של האומד ! ולא השונות של האומד ! כלומר בכל מקום שכתוב θ נשנה ל$\hat{\theta}$ בהתאמה ! .
<u>פילוסופיה על $\hat{V}(\theta)$</u> יכולתי לבחור איזה אומד לשונות של האומד שאני רוצה ולכאורה הייתי צריך לבחור בS_n אבל במקרים בהם הצגתי את $\theta = \hat{\theta}$ **וניתן לתאר** את השונות בעזרת θ **אז אני אתאר** את השונות בעזרת $\hat{\theta}$.
כאשר הסיבה לכך היא ביצועית ! כלומר ההתכנסות היא יותר מהירה!
דוגמא אם $X_1, \dots X_n \overset{i.i.d}{\sim} Exp(\lambda)$ ו $\overline{X_n} = \frac{1}{\lambda}$ אזי $\hat{V}(\overline{X_n}) = \frac{1}{(\frac{1}{\overline{X_n}})^2}$

1) $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \hat{V}_{X_1}(\theta))$
2) תהיי $g : H \to \mathbb{R}$ גזירה ברציפות אזי

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} N(0, \hat{V}(\theta) \cdot [g^{'}(\theta)]^2)$$

דוגמא:

תחת תנאי המשפט הגבול המרכזי $\sqrt{n}(\bar{X_n}-\mu)\xrightarrow{d}N(0,\sigma^2)$
הערה הקצב α_n היא סדרה יחידה !
מתרגול 9
שיעור 18

5.2 רווחי סמך מדוקים ומקורבים

כאן מדובר על אמידה בעזרת רווחי סמך !

הגדרה 44.2 רווח סמך מדויק
אינטרוול סטטיסטי $I(X)=[a_-(X),a_+(X)]$ הוא רווח סמך ברמת ביטחון הסתברותית $1-\alpha$
כלומר $\forall\theta\in\Theta$ מתקיים כי $P_\theta(\theta\in I(X))\geq 1-\alpha$
משתמשים בדר"כ במשתנה פיבוט שזה משתנה שהתפלגותו לא תלויה בθ והוא נוצר מ$X_1,\dots X_n$כלומר פונקציה של מרחב המדגם שלי.
בד"כ לא ניתן לבנות רווח סמך מדויק הינה מספר מקרים שכן אפשר לבנות רווח סמך מדויק !

1. אם $N(\theta,1)\overset{i.i.d}{\sim}X_1,\dots X_n$ אזי ניתן לבנות פיבוט $\sqrt{n}\cdot\frac{|\bar{X_n}-\theta|}{1}\sim N(0,1)$ ואז נקבל $P(\sqrt{n}\cdot\frac{|\bar{X_n}-\theta|}{1}\leq Z_{1-\frac{\alpha}{2}})=2\Phi(Z_{1-\frac{\alpha}{2}})-1$
2. אם $X_1,\dots X_n\sim U(0,\theta)$ אזי ניתן לבנות פיבוט $\frac{X_1}{\theta}\sim U(0,1)$

רווחי סמך מקורבים
יהי $\hat{\theta}_n$ סדרת אומדים כלשהו ל θ שמבוססת על המדגם $X_1\dots X_n$
כך ש ו $E(\hat{\theta}_n^2)<\infty$
בנוסף $\hat{\theta}_n\xrightarrow{p}\theta$סדרה עקיבה ל$\theta$
כאשר אסימפטוטית $\hat{\theta}_n$ מתפלגת נורמלית ! עם אומד לשונות של האומד $V(\hat{\theta})$
אזי לפי משפט הגבול המרכזי מתקיים כי $\sqrt{n}(\hat{\theta}_n-\theta)\xrightarrow{d}N(0,var(\hat{\theta}))$.
על סמך זה נבנה את רווח הסמך $1-\alpha=P(\hat{\theta}-c\leq\theta\leq\hat{\theta}+c)=P(-c\leq\hat{\theta}-\theta\leq c)=P(\sqrt{n}\cdot\frac{-C}{\sqrt{V(\hat{\theta})}}\leq\sqrt{n}\cdot\frac{\hat{\theta}-\theta}{\sqrt{V(\hat{\theta})}}\leq\sqrt{n}\cdot\frac{C}{\sqrt{V(\hat{\theta})}})$
לכן אסיק כי $Z_{1-\frac{\alpha}{2}}=\sqrt{n}\cdot\frac{C}{\sqrt{V(\hat{\theta})}}$ נבודד את C ונציב בביטוי $P(\hat{\theta}-c\leq\theta\leq\hat{\theta}+c)$ ונקבל

$$C=\hat{\theta}_n\pm\frac{1}{\sqrt{n}}\sqrt{\hat{V}(\theta)}\cdot Z_{1-\frac{\alpha}{2}}$$

בדר"כ נהוג להשתמש ב$\hat{\theta}=\bar{X_n}$ ואז נקבל$C=\bar{X_n}\pm\frac{1}{\sqrt{n}}\sqrt{\hat{V}(\theta)}\cdot Z_{1-\frac{\alpha}{2}}$

רווח סמך על פי שיטת $\triangle$
יהי $\hat{\theta}_n$ סדרת אומדים כלשהו ל θ שמבוססת על המדגם $X_1\dots X_n$כך ש $E(\hat{\theta}_n^2)<\infty$
בנוסף $\theta\xrightarrow{P}\hat{\theta}$
כאשר אסימפטוטית $\hat{\theta}_n$ מתפלגת נורמלית ! עם האומד לשונות של האומד $\hat{V}(\theta)$
כלומר אם $\sqrt{n}(\hat{\theta}_n-\theta)\xrightarrow{d}N(0,\sigma^2)$ כאשר σ^2 זה השונות של דגימה בודדת בסדרה $\hat{\theta}_n$!
ותהיי $g:\Theta\longrightarrow\mathbb{R}$ פונק גזירה ברציפות לפי שיטת הדלתא מתקיים כי $\sqrt{n}(g(\hat{\theta})-g(\theta))\xrightarrow{d}N(0,[g'(\theta)]^2\cdot V(\theta))$
אם נרצה לבנות רווח סמך ל$g(\theta)$ ברמת מובהקות $1-\alpha$ נקבל
כעת עשינו רדוקציה לבעיה הקודמת !
כלומר נשים לב כי $[g'(\theta)]^2\cdot V(\theta)=\sigma(\theta)$ וכשנבנה רווח סמך נציב את $g(\hat{\theta})$ בתוך $\sigma(\theta)$ האומד לשונות של האומד ! ואז נבנה את רווח הסמך
נשים לב שאם $V(\theta)>0$ וקיימת פונק g גזירה ברציפות כך ש$[g'(\theta)]^2\cdot V(\theta)=L>0$
ואז רווח סמך לθ ברמת סמך $1-\alpha$ הינו

$$C=g^{-1}(g(\hat{\theta})\pm\frac{1}{\sqrt{n}}\cdot Z_{1-\frac{\alpha}{2}}L)$$

משבוע 10 אצל פבל !
יעילות אסימפטוטית רווח סמך לאומד נראות מירבית
מהי סדרת האומדים האופטימלית $\hat{\theta}_n$זו שאני צריך n הקטן ביותר על למנת להגיע להתפלגות נורמלית !
מסתדר שאומד בייס ואומד נראות מירבית הם הטובים ביותר !
נאמר ש$\hat{\theta}^*$ הוא אומד יעיל אסימפטוטית אם $\theta\in\Theta$ $E_\theta(\sqrt{n}(\hat{\theta^*}-\theta)^2)=\frac{1}{I_{\vec{X}}(\theta)}$
הסיבה לכך נובעת מחומר שלא במסגרת הקורס הזה.
תהיי $L(X,\theta)=\prod_{i=1}^{n}f(X_i,\theta)$ פונקצית נראות מירבית
בנוסף $I(\theta)=E_\theta(\frac{\partial logf(X_1,\theta)}{\partial\theta})<\infty$
ו $\hat{\theta}=argmax_\theta L(X,\theta)$
אזי נקבל כי $\sqrt{n}(\hat{\theta}-\theta)\xrightarrow{d}N(0,\frac{1}{I_{\vec{X}}(\theta)})$ לכל $\theta\in\Theta$
משבוע 10 אצל פבל
יש הנחות נוספות למשפט שצריכות להתקיים שפבל לא הסביר

טענה 45.2 רווחי סמך מקורבים להתפלגיות מוכרות

1. תהיי $X_1,\dots X_n \overset{i.i.d}{\sim} Exp(\lambda)$ אזי $[\frac{1}{\overline{X_n}} - \sqrt{\frac{1}{n\cdot(\overline{X_n})^2}}\cdot Z_{1-\frac{\alpha}{2}}, \frac{1}{\overline{X_n}} + \sqrt{\frac{1}{n\cdot(\overline{X_n})^2}}\cdot Z_{1-\frac{\alpha}{2}}]$

2. תהיי $X_1,\dots X_n \overset{i.i.d}{\sim} ber(\theta)$ אזי $[\overline{X_n} - \sqrt{n[\overline{X_n}(1-\overline{X_n})]}\cdot Z_{1-\frac{\alpha}{2}}, \overline{X_n} + \sqrt{n[\overline{X_n}(1-\overline{X_n})]}\cdot Z_{1-\frac{\alpha}{2}}]$

3. תהיי $X_1,\dots X_n \overset{i.i.d}{\sim} Po(\theta)$ אזי $[\overline{X_n} - \sqrt{n\overline{X_n}}\cdot Z_{1-\frac{\alpha}{2}}, \overline{X_n} + \sqrt{n\overline{X_n}}\cdot Z_{1-\frac{\alpha}{2}}]$

4. תהיי $X_1,\dots X_n \overset{i.i.d}{\sim} N(\mu,\sigma^2)$ עם σ^2 ידועה רווח סמך לμ אזי $\left[\overline{X} - z_{1-\alpha/2}\sigma/\sqrt{n}, \overline{X} + z_{1-\alpha/2}\sigma/\sqrt{n}\right]$

5. תהיי $X_1,\dots X_n \overset{i.i.d}{\sim} N(\mu,\sigma^2)$ עם שונות לא ידועה אזי $[\overline{X}_n - \frac{S_n}{\sqrt{n}}t_{n-1,1-\alpha/2}, \overline{X}_n + \frac{S_n}{\sqrt{n}}t_{n-1,1-\alpha/2}]$

6. תהיי $X_1,\dots X_n \overset{i.i.d}{\sim} N(\mu,\sigma^2)$ רווח סמך לσ^2 אזי $[\frac{(n-1)S_n^2}{\chi^2_{n-1,\frac{\alpha}{2}}}, \frac{(n-1)S_n^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}]$

6.2 בדיקת השערות

$X \sim P_{\theta\in\Theta}$ כאשר $\theta \in \mathbb{R}^d$ מחלקים את מרחב הפרמטרים שלנו Θ ל 2 מרחבים $\Theta_0 \oplus \Theta_1 = \Theta$
כאשר $H_0 : \theta \in \Theta_0$
$H_1 : \theta \in \Theta_1$
נגדיר מבחן לדחיית השערת ה0 $\delta : X \to \{0,1\}$ כאשר $\delta = 1_{\{T(x)\geq C\}}$ כאשר $T(x)$ נקרא סטטיסטי מבחן !(פיבוט)
קשר בין בדיקת השערות לרווח סמך
יש קשר בין רווח סמך לבין בדיקת השערות אבל נשים לב שהמטרה שלי בבדיקת השערות היא שT יהיה פיבוט בלבד !
ולא בהכרח ש$E(T)=\theta$ למשל עבור $X_1,\dots X_n \sim Exp(\lambda)$ פיבוט טוב כי הוא $\sqrt{n}\cdot\frac{\bar{X}-\frac{1}{\lambda}}{\sqrt{\frac{1}{\lambda^2}}}\dot\sim N(0,1)$
אבל על מנת לעשות רווח סמך לλ אין ברירה ונאלץ לעבור לדרך שיטת ה$\triangle$!
נשים לב שאם בניתי רווח סמך סביב θ ברמת סמך של $1-\alpha$ אם תחת H_0 θ לא נמצא באינטרוול נדחה את השערת ה0 אחרת לא נדחה את השערת ה0
אם $\delta = 1$ כלומר $T(X)\geq C$ נדחה את השערת ה0
אם $\delta = 0$ כלומר $T(X) < C$ לא נדחה את השערת ה0 !
כאשר C הוא מספר קבוע והוא הערך הקריטי ללדחות את השערת ה0 או לא !
עוצמות וטעיות
טעות מסוג I דחיתי את השערת ה0 כאשר היא נכונה !
טעות מסוג II קיבלתי את השערת ה0 כאשר היא לא נכונה!
אסוציאציה בדיקת השערות האם אדם חולה
טעות מסוג I אמרתי שאדם לא חולה כשבפועל הוא כן חולה (הוא יכול למות בגללי)
טעות מסוג II אמרתי שחולה כשבפועל הוא לא חולה ! (טיפלתי בו סתם)
טעות מסוג I היא החמורה יותר וממנה אני רוצה להימנע ככל שאפשר !בדר"כ נדרוש $P_\theta(T(X)\geq c)\leq 0.05$
נגדיר פונקציית עוצמה
$\theta\in\Theta$ $\pi(\theta,\delta)=E_\theta(\delta)=P_\theta(T(X)\geq c)$ היא מכילה בתוכה את החישוב של 2 הטעיות
$\pi(\theta_0,\delta)=P_{\theta\in\Theta_0}(T(X)\geq c)$ רמת המובהקות לדחות את השערת ה0
$\pi(\theta_1,\delta)=P_{\theta\in\Theta_1}(T(X)\geq c)$ עוצמת המבחן לדחות את השערת ה0
$P_{\theta\in\Theta_0}(T(X)\geq c)$ פונק העוצמה מבטאת טעות מסוג I דחיית H_0 כאשר היא נכונה
עוצמת המבחן
$P_{\theta\in\Theta_1}(T(X)\geq c)$ פונק העוצמה מבטאת טעות מסוג II דחיית H_0 כאשר H_1 נכונה
במילים של מיכה $\pi = P_{H_1}(reject\ H_0)$
$\pi(\theta_1,\delta)$ זה חישוב עוצמת המבחן
לרוב נשתמש ברווחי סמך מקורבים לבדיקת השערות ונקבל C_α
קרי $P_\theta(T(X)\geq c_\alpha)=0.05$
לבנות פיבוט
ו$T(X)$ נרצה שהוא יהיה פיבוט כלומר משתנה שמה$data$ שלי התפלגותו לא תלויה בθ
לרוב נתחיל האומד נקודתי ונפתח אותו לפיבוט המתאים

1. $P_{H_0}(\bar{X}>C)=0.05$ ומכאן נבנה את הפיבוט המתאים לבדיקת השערות עבור התוחלת.

2. $P_{H_0}(S_n^2>C)=0.05$ ומכאן נבנה את הפיבוט המתאים לבדיקת השערות עבור השונות σ^2

3. נשים לב שזה לא תמיד המצב תוך שימוש בלמה של ניצן פירסון מקבלים פיבוטים שונים !

המשחק פה הוא למצוא את c_α בבדיקת השערות !.
המון פעמים פרקטית נעשה מהלך כזה $P_{H_0}(T(x)\leq C)=1-\alpha \iff P_{H_0}(T(x)>C)=\alpha$
ערך p
$p-value$
נגדיר $p^{val}=p(t)=P_{\theta\in\Theta_0}(T(x)\geq t)$ כאשר $t=T(x)$ במילים אחרות $p^{val}=\underset{t\in\mathbb{R}}{inf}P_{H_0}(T(x)\geq t)$
$T(X)$ סטטיסטי תחת H_0
t מה שהתקבל במדגם שלי
הוא ה
ערך P מוגדר כהסתברות לקבל תוצאה זהה לזו שהתקבלה מסטטיסטיקת המדגם או תוצאה "קיצונית" ממנה,

תחת ההנחה שהמדגם התקבל מהתפלגות מסוימת, הנחה זו מכונה השערת האפס.
טרם ביצוע המבחן נקבעת רמת המובהקות)או רמת הסמך(של המבחן, אשר מסומנת באות α.
אם $\alpha < p^{val}$ דוחים את השערת ה0
אם $\alpha > p^{val}$ לא דוחים את השערת ה0
אם הערך P קטן או שווה ל-α)רמת מובהקות המבחן שנקבעה זה מרמז כי התוצאות שקיבלנו מהמדגם אינן מתיישבות עם ההנחה כי השערת האפס היא נכונה ולכן יש לדחות את אותה ההשערה.

1.6.2 בדיקת השערות התפלגיות מוכרות -מיכה

טענה 46.2 בדיקת השערות עבור **התפלגות נורמלית לשונות עבור** $N(\mu,\sigma^2)$
כאשר ההשערה היא $H_0:\sigma^2=13$ ו $H_1:\sigma^2\neq 13$(סתם דוגמא)
אחשב מהו $P\left(\chi^2_{n-1,\frac{\alpha}{2}}\leq\frac{\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}{\sigma^2}\leq\chi^2_{n-1,1\frac{\alpha}{2}}\right)=P\left(\frac{(n-1)S_n^2}{\chi^2_{n-1,1\frac{\alpha}{2}}}\leq\sigma^2\leq\frac{(n-1)S_n^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right)=1-\alpha$
משתנה שלא תלוי בשונות ובתוחלת $(n-1)S_n^2=\sum_{i=1}^n(x_i-\bar{x})^2$

$$\left[\frac{(n-1)S_n^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}},\frac{(n-1)S_n^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$$

נדחה את השערת ה0 אם σ^2 לא שייך לרווח הסמך
לא נדחה את השערת ה0 אם σ^2 שייך למקטע

טענה 47.2 בדיקת לתוחלת עבור **התפלגות נורמלית כאשר התוחלת והשונות לא ידועים**
$H_0:\mu=\mu_0$
$H_1:\mu>\mu_0$
$\mathbb{P}_{H_0}(\bar{X}>C)=\mathbb{P}_{H_0}(\frac{\overline{X}-\mu_0}{S_n/\sqrt{n}}>\frac{c-\mu_0}{\hat{\sigma^2}/\sqrt{n}})=\mathbb{P}_{H_0}(T>\frac{c-\mu_0}{\hat{\sigma^2}/\sqrt{n}})=\alpha$
כאשר $S_n=\hat{\sigma}^2=\frac{1}{n-1}\sum_{i=1}^n\left(x_i-\overline{X}\right)^2$
כי כך בדיוק נראה משתנה $T=\frac{\overline{X}-\mu}{S_n/\sqrt{n}}\sim t_{(n-1)}$

טענה 48.2 בדיקת השערות על תוחלת **בהתפלגות נורמלית ושונות ידועה**
$X_1,\ldots X_n\sim N(\mu,\sigma^2)$
$H_1:\mu=\mu_0\ H_0:\mu=\mu_1\ \ *(\mu_1<\mu_0)$
$\mathbb{P}_{H_0}(\bar{X}\overset{*}{\leq}C)=\mathbb{P}_{H_0}(\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}}<\frac{c-\mu_0}{\sigma/\sqrt{n}})=\Phi\left(\frac{C-\mu_0}{\sigma/\sqrt{n}}\right)$
ואקבל כי $\frac{C-\mu_0}{\sigma/\sqrt{n}}=-z_{1-\alpha}$
לאחר העברת אגפים דחה את השערת ה0
אם $\overline{X}\leq\mu_0-z_{1-\alpha}\sigma/\sqrt{n}$

טענה 49.2 קביעת גודל המדגם במקרה של פרופרציית תמיכה
בודקים מהו $\mathbb{P}_{H_0}(C_1\leq F\leq C_2)=1-\alpha$
ומחשבים מהם C_2-C_1
טיפ מתקיים כי $Max[\hat{P}(1-\hat{P})]=0.5$

2.6.2 הלמה של נימן פירסון

טענה 50.2 מבין כל המבחנים לבדיקת השערות פשוטות נרצה למצוא מבחן שהוא בעל העוצמה הmax
נניח כי $L(X,\theta)$ קיימת אזי מבחן יחס הנראות הינו

$$T(X)=\frac{L(X,\theta_1)}{L(X,\theta_0)}$$

עם רמת מובהקות α !
נשים לב שהמבחן הוא תמיד חד צדדי !! בלי ערך מוחלט על הסטטיסטי!
שיעור 19 שבוע 11 אצל פבל !
ניסוח שקול:
מבחן UMP לבדיקת השערות פשוטות הוא מבחן יחס נראות מירבית.
בפועל מראים בשיעור 20 ש
מבחן **נימן פירסון טוב גם לבדיקת השערות מורכבות ! כלומר הוא מבחן** UMP
תחת התנאים הבאים:

1. נגדיר $\delta^*(X)=1_{\{T(x)\geq C\}}$ ואם $\pi(\delta^*(x))=E\delta^*$ זו פונקציית העוצמה היא פונקציה מונ עולה חלש ביחס ל θ !
2. $\delta^*(X)$ קבוע ביחס לθ_1

טענה 51.2 אופטימליות בין מבחנים וקצת פילוסופיה
נרצה לענות על השאלה בהינתן 2 מבחנים לבדיקת השערות איזה מבחן נעדיף?
תחילה נשווה רק בין 2 מבחנים בעלי אותה רמת מובהקות קרי α אחרת לא עשינו בזה כלום !
בהינתן בדיקת השערות לא פשוטה קרי כלומר H_0 יותר גדול מנק בודדת קרי $H_0\oplus H_1=\Theta\subset\mathbb{R}^d$.

נגדיר רמת מובהקות בהשערות לא פשוטות בצורה הבאה $\underset{\theta\in\Theta_0}{Sup} P_{\theta_0}(T(X)\geq C)=\alpha$.

יכולנו להגדיר בצורה שונה (אופטימליות בייס-לא נלמד-).

השוואה בין מבחנים

נאמר שמבחן $\delta_2(X)$ הוא עצמתי יותר ממבחן $\delta_1(X)$ אם

$\pi_{|H_1}(\delta_1(X))=E_\theta(\delta_1(X))\geq E_\theta(\delta_1(X))=\pi_{|H_1}(\delta_2(X))\ \forall\theta\in H_1$

וקיים $\theta\in H_1$ (קיים לפחות אחד יש מצב שיהיו יותר !)כך שהא"ש הופך לחריף !.

לרוב לא יהיה ניתן להשוואת בין 2 מבחנים כמו שראינו על פונקציות סיכון אמידה.

מבחן UMP

מבחן $\delta^*(X)$ הוא UMP

אם לכל $\delta(X)$ מבחן כלשהו מתקיים כי $\forall\theta\in H_1$ $E_\theta\delta^*(X)\geq E_\theta\delta(X)$.

כלומר δ^* הוא המבחן עם העוצמה הגדולה ביותר ! ברמת מובהקות α

שיעור20

שבוע 11

3.6.2 בדיקת השערות מורכבות

טענה 52.2 תחת תנאים מסוימים ניתן הלמה של נימן פרסון טובה גם להשערה מורכבת

כפי שאנו רואים בטענה כאן

נניח מודל סטטיסטי $P_{\theta\in H}$ כאשר $\Theta\in\mathbb{R}$ ויהי δ^* מבחן $F-N$ יחס נראות מירבית ברמת מובהקות α

לבדיקת השערות פשוטה קרי

$H_0:\theta=\theta_0$

$H_1:\theta=\theta_1$

כאשר $\theta_1>\theta_0$

אם

1. δ^* קבועה ביחס לθ_1
2. $\pi(\theta,\delta^*)=E_\theta(\delta^*)$ היא פונקציה מונ עולה חלש ביחס לθ

אזי δ^* הוא מבחן UMP גם לבדיקת השערות מהצורה

$H_0:\theta<\theta_0$

$H_1:\theta\geq\theta_1$

מתרגול 12 ומשיעור 20

טענה 53.2 מבחן $GLRT$

יהי δ^* מבחן יחס נראות בעל רמת מובהקות α

מבחן $GLRT$ **לבדיקת השערות מורכבות** דוגמא גנרית לבדיקת מבחנים.

לא תמיד ניתן להשתמש בלמה של נימן פרסון לבדיקת השערות מורכבות

לכן נרצה שתהיה לנו שיטה גנרית למציאת מבחנים.

היא לא השיטה הכי טובה שיש אבל יש בא הגיון.

נשים לב שהיא מקרה מוכלל של נימן פירסון

$$N(X)=\frac{\underset{\theta\in H_1}{Sup}L(X,\theta)}{\underset{\theta\in H_0}{Sup}L(X,\theta)}\overset{*}{=}\frac{L(\hat\theta^{(1)},X)}{L(\hat\theta^{(0)},X)}\overset{**}{=}\frac{\underset{\theta\in H}{Sup}L(X,\theta)}{\underset{\theta\in H_0}{Sup}L(X,\theta)}$$

* כאשר הSup מושג ! $\hat\theta^{(j)}=argmax_{\theta\in H_j}L(X,\theta)$

** אם פונק ה$L(X,\theta)$רציפה. נשים לב כי $\underset{\theta\in H}{Sup}L(X,\theta)$ לא מאולץ אבל $\underset{\theta\in H_0}{Sup}L(X,\theta)$ מותנה תחת השערת ה0 !

אז $\delta(X)=1_{\{N(X)>C\}}$

שיעור 20 שבוע 12

טענה 54.2 מבחן $wild$ (המקרה הרב מימדיי) לבדיקת השערות מורכבות

$H_0:\theta\in\Theta_0$

$H_1\theta\in\Theta_1$

כאשר $\Theta=\Theta_0\oplus\Theta_1\subset\mathbb{R}^k$

נציג את H_0 בעזרת פונקציה כלשהי R כאשר $H_0=\{\theta\in H\mid R(\theta)=0\}$ ו $R:H\to\mathbb{R}^r$ $1\leq k\leq r$

נניח כי $L(\theta,X^n)\longrightarrow\theta$ מקיימת את התנאים המבטיחים

וגם $\hat\theta$הינו אומד נראות מירבי כשהוא מקיים $\sqrt{n}(\hat\theta-\theta)\xrightarrow{d}N(0,I(\theta)^{-1})$

אזי מתקיים כי

$$2log(\hat\theta)\xrightarrow{d}\chi^2_r\ \forall\theta\in H_0$$

נהוג לומר כי סטטיסטי הוא נטול התפלגות אם ההתפלגות שלו לא תלויה בθ !

מבחן $wild$ ווילדס (המקרה החד מימדיי)

יהי $X_1,\dots X_n$מדגם מצפיפות $f(X,\theta)$ ויהי $\frac{L(\hat{\theta}^{(1)},X)}{L(\hat{\theta}^{(0)},X)} \overset{*}{=} \frac{\underset{\theta\in H_1}{Sup}\ L(X,\theta)}{\underset{\theta\in H_0}{Sup}\ L(X,\theta)} = \lambda(X)$ סטטיסטי $GLRT$ לבדיקת השערות מורכבות קרי

$H_o:\theta=\theta_0$
$H_1:\theta\neq\theta_0$

כאשר θ_0היא נק פנימית באינטרוול $H\subset\mathbb{R}$ אזי **תחת** H_0 **מתקיים כי**

$$2log\lambda_n(X) \underset{n\to\infty}{\overset{d}{\to}} \chi_1^2$$

מתרגול 12

שיעור 21 (המקרה הרב מימדיי)

3 דגמאות נגדיות חשובות

טענה 1.3 גם אם המודל אינו ניתן לזיהוי זה לא אומר שלא ניתן להעריך את θ !

דוגמא נביט במודל הרגרסיה הבא

$$Y_j = < x,\beta > + \varepsilon_j \ \ j=1\dots,n \ \ \beta\in\mathbb{R}^p$$

עבור הוקטור המתוקן $x\in\mathbb{R}^p$ נשים לב כי ε_jלא ניתן לזיהוי עבור $\beta\in\mathbb{R}^p$

אבל ניתן להעריך את $< x,\beta >$!

שבוע 2

טענה 2.3 קיימת דוגמא כך שלא קיים לה אומר חסר הטייה

אם $ber(\theta) \overset{i.i.d}{\sim} X_1,\dots X_n$ נגדיר $q(\theta)=\frac{\theta}{1-\theta}$ אזי לא קיים אומד חסר הטייה ל$q(\theta)$!

שבוע 6

$$f\begin{pmatrix}y\\z\end{pmatrix} = E[X|\begin{pmatrix}Y\\Z\end{pmatrix}=\begin{pmatrix}y\\z\end{pmatrix}] = \sum_{x\in Supp(X)} x\cdot P(X|\begin{pmatrix}Y\\Z\end{pmatrix}=\begin{pmatrix}y\\z\end{pmatrix}) = \sum_{x\in Supp(X)} x\cdot\frac{P(X=x,Y=y,Z=z)}{P[\begin{pmatrix}Y\\Z\end{pmatrix}=\begin{pmatrix}y\\z\end{pmatrix}]}$$