משפטים והגדרות מודלים ב

עמית ירון

9 באפריל 2022

1 מבוא לסטטיטיקה א-פרמטרית :

מוטיבציה:
אז בסטטיטיקה א פרמטרית אנחנו עומדים פונקצית התפלגות ולא פרמטר בודד בגלל זה $dim[\Theta]=\infty$
ישנם 3 בעיות מרכזיות בסטטיטיקה

1. אמידה נקודתית של f
2. רווח סמך על f
3. בדיקת השערות f

מדדים ונורמות מוכרות :

1. $MSE(\hat{f},f)=E_F[\hat{f}-f]^2=var(\hat{f})+[\hat{f}-E(\hat{f})]^2$ נקרא גם $Mean\ Squared\ Error$
   (א) MSE **הוא עבור נקודה קבועה** x_0**! או עבור וקטור קבוע** $\overrightarrow{x_0}$
2. $||\hat{f}-f||_\infty=\underset{n\to\infty}{lim}\underset{x\in\mathbb{R}}{Sup}|\hat{f}_n-f|$ התכנסות במידה שווה $uniform\,conversion$ או $Mean\ Uniform\ Error$
   (א) **מדד לכל** $x\in\mathbb{R}$ **רצף של נקודות אמידה פונקציונלית**
3. נורמת L^2 הינה $||\hat{f}-f||_{L2}=E_F[\int\limits_{\mathbb{R}}(\hat{f}-f)^2dFx]=E_F[\int\limits_{\mathbb{R}}(\hat{f}-f)^2\cdot f_xdx]$
4. **נורמת** $MISE$ **הינה :** $MISE=E_F\int\limits_{\mathbb{R}}(\hat{f}-f)^2dx=$ ההבדל מנורמה L^2 היא שבאינטגרל יש dx ואין $dF(x)$! $MeanUniformError)$
5. $||\sqrt{n}(\hat{F}(X^k)-F(X^k))||_{\Sigma_F}=\sqrt{n}(\hat{F}(X^k)-F(X^k))^T\cdot\Sigma_F^{-1/2}\cdot\Sigma_F^{-1/2}\cdot\sqrt{n}(\hat{F}(X^k)-F(X^k))$
6. ראה פרק 4אצל פבל

1.0.1 טריקים חשובים שהופיעו במהלך הקורס

1. וקטור מקרי $\overrightarrow{X}=\begin{pmatrix}x_1\\\vdots\\x_k\end{pmatrix}$מתנהג מולטינומי כלומר ($\overrightarrow{X}\sim multi(n,\begin{pmatrix}p_1\\\vdots\\p_k\end{pmatrix}$ אם כל רכיב בוקטור מתנהג בינומי כלומר אם $x_i\sim Bin(n,p_i)\,\forall i\in$ $[k]$ מתרגיל 1 בשיעורי בית !
2. משפט הגבול המרכזי אם $f\overset{i.i.d}{\sim}X_1,...X_n$ וגם $EX_1<\infty,EX_1^2<\infty$ אז $\begin{cases}\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\xrightarrow{d}N(0,1)\\\Updownarrow\\\sqrt{n}\cdot(\overline{X}-\mu)\xrightarrow{d}N(0,\sigma^2)\end{cases}$ במקרה שלנו $\overline{X}=\hat{F}_n(x_0)$!
3. אם יש לנו ביטוי $\underset{n\to\infty}{lim}E(Y_n)$ ונרצה להחליף אותו ב$E(\underset{n\to\infty}{lim}Y_n)$
   יש לנו קריטריון שאומר כי Y_n אינטגרבילית במידה שווה $\Longleftrightarrow$ קיימת ϕ כך ש $\begin{cases}\underset{x\to\infty}{lim}\frac{\phi(x)}{x}=\infty\\\underset{n\in\mathbb{N}}{Sup}E\phi(|Y_n|)<\infty\end{cases}$ **מתרגיל 1 שאלה 7**
4. נגזרת כיוונית מוגדרת להיות $\underset{t\to0}{lim}\frac{f(\overrightarrow{x}+t\frac{\overrightarrow{v}}{||\overrightarrow{v}||})-f(\overrightarrow{x})}{t}$
   כאשר אנו מחשבים את $\overrightarrow{x}$בכיון $\overrightarrow{v}$ בנוסף אם F היא דיפרנצאבילית $D_{\overrightarrow{v}}f=<\nabla f,\overrightarrow{v}>$מויקיפדיה
5. $1_{\{X_j<b\}}-1_{\{X_j<a\}}=1_{\{a<X_j\leq b\}}$ שימו לב לא"ש החזק בצד שמאל !

6. $\int\limits_{\mathbb{R}} 1_{x\in[a,b]}dF = F(b) - F(a)$

7. אם $X_1,...X_n \overset{i.i.d}{\sim} N(\mu,\sigma^2)$ ואומדים את $\theta = \begin{pmatrix}\mu\\ \sigma^2\end{pmatrix}$ בעזרת MLE שזה יוצא $\hat{\theta} = \begin{pmatrix}\overline{X}\\ S_n^2\end{pmatrix}$
נקבל כי $\sqrt{n}[\hat{\theta}-\theta] \xrightarrow{d} N(0,\begin{pmatrix}\sigma^2 & 0\\ 0 & 2\sigma^4\end{pmatrix})$ כי $\begin{cases} Var(\sqrt{n}\overline{X}) = \sigma^2 \\ Var(\sqrt{n}S_n^2) = 2\sigma^4\end{cases}$

8. חישוב אחוזון של התפלגות נורמלית **תרגיל 2 שאלה 3 אצל פבל** $q(p) = \mu + \sqrt{\sigma^2}\cdot Z_p$

9. **נוסחת בינום של ניוטון:** $(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y^1 + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n-1}x^1y^{n-1} + \binom{n}{n}y^n$
   (א) $(x+y)^7 = x^7 + 7x^6y + 21x^5y^2 + 35x^4y^3 + 35x^3y^4 + 21x^2y^5 + 7xy^6 + y^7$
   (ב) $(x+y)^6 = x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6$
   (ג) $(x+y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5$
   (ד) $(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$
   (ה) $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$

10. **ממבחן מועד א של פבל:**
   (א) אם $X_1,...X_n \overset{i.i.d}{\sim} F$ וF לא ידועה $E_F[F(\overline{X}_n)] = F(\overline{X}_n) = F(\overline{X_{n-1}})$

1.0.2 אמידה נקודתית של f על סמך פונק ההתפלגות האמפירית

אנחנו מניחים כי $f \overset{i.i.d}{\sim} X_1,...X_n$ כאשר f היא לא ידועה אבל $c.d.f$ הפונק הצוברת שלה מקיימת את התכונות שידועות לנו מהסתברות :

1. פונק לא יורדת
2. $f:\mathbb{R}\to[0,1]$
3. $\lim\limits_{x\to\infty} F(x) = 1$
4. $\lim\limits_{x\to-\infty} F(x) = 0$

הצעה לאומדן נקודתי של f הינה היסטוגרמה כלומר פונק ההתפלגות האמפירית $\hat{F}_n(x) = \frac{1}{n}\cdot\sum\limits_{j=1}^{n} 1_{\{X_j\le x\}}$

טענות בסיסיות לפונק ההתפגלות האמפירית :

1. **מדובר באומד חסר הטייה** כלומר $E_F[\hat{F}_n(x)] = F(x)\ \forall x\in\mathbb{R}$
2. **לפי מדד** MSE **עבור** $x\in\mathbb{R}$ **נקודה קבועה** נקבל כי

$$MSE(\hat{F}_n(x),F_n(x)) = E_F[\hat{F}_n(x)-F_n(x)]^2 = Var[\hat{F}_n(x)] + \underbrace{Bias(\hat{F}_n(x),F_n(x))}_{=0} = F(x)[1-F(x)] \le \frac{1}{n}\cdot\frac{1}{4}$$

   (א) חישוב מומנט שני של $n\cdot\hat{F}_n(x)\sim Bin(n,F(x))$ $E_F[\sqrt{n}[\hat{F}_n(x)-F_n(x)]]^2 = Var[n\cdot[\hat{F}_n(x)-F_n(x)]] = F(x)(1-F(x)) \le \frac{1}{4}$
   (ב) $Var(\hat{F}_n) = \frac{1}{n}F(x)[1-F(x)]$ מהשיעורי בית
   (ג) $F(x)(1-F(x)) \le F(x) + (1-F(x)) = 1$ מהתרגיל ה1
   (ד)

$$\int\limits_{\mathbb{R}} F(x)(1-F(x))dx = \int\limits_{-\infty}^{0} F(x)(1-F(x))dx + \int\limits_{0}^{\infty} F(x)(1-F(x))dx \le \int\limits_{-\infty}^{0} F(x)dx + \int\limits_{0}^{\infty}(1-F(x))dx = EX^- + EX^+ = E|X|$$

מתרגיל 1

   (ה) $E_F||\sqrt{n}(F_n-F)||_2^2 \le E_FX^+ + E_FX^- = E_F|X|$ מהתרגיל ה1
   (ו) מסקנה רלוונטית היא ש $\begin{cases} = E\int\limits_{\mathbb{R}}(\hat{F}(x0)-F(x0))^2dF \iff \hat{F}_n(x)\xrightarrow{L^2} F \\ \lim\limits_{n\to\infty} Sup|\hat{F}(x0)-F(x0)| \longrightarrow 0 \iff \hat{F}_n(x) \overset{Uniform}{\longrightarrow} F \\ \hat{F}_n(x)\xrightarrow{p} F \qquad\qquad LLM\ Rule\end{cases}$

3. כאשר F היא בדידה עם אטומים $\begin{cases}\{x_1,...x_k\}\\ \{p_1,...p_k\}\end{cases}$ אז

(א) $D_n=\underset{1\le j\le k}{Max}|\hat{p}_j-p_j|$ כאשר $\hat{F}_n(x_j)=\frac{1}{n}\sum\limits_{i=1}^{n}1_{\{X_i=x_j\}}=\hat{p}_j$

(ב) $\sqrt{n}[\hat{p}-p]\xrightarrow{d}N(0,\Sigma_p)=Y$ כאשר $\Sigma_p=dig(p)-p^Tp$ לכן $D_n\xrightarrow{d}\underset{1\le j\le k}{Max}Y$ בנוסף $p=\begin{pmatrix}\hat{F}_n(x_1)\\ \vdots\\ \vdots\\ \hat{F}_n(x_k)\end{pmatrix}$

4. **מקרה מעניין של פונקציונל לינארי** $\begin{cases}T(\hat{F}_n)=\int\limits_{\mathbb{R}^d}\psi d[\hat{F}_n]=\frac{1}{n}\sum\limits_{j=1}^{n}\psi(X_j)=\overline{\psi}_n\\ T(\hat{F}_n)=\int\limits_a^b\psi d[\hat{F}_n]=\frac{1}{n}\sum\limits_{j=1}^{n}\psi(X_j)\cdot 1_{\{a\le\psi(x_i)\le b\}}\\ \int\limits_{\mathbb{R}}1_{x\in[a,b]}d\hat{F}_n=\hat{F}_n(b)-\hat{F}_n(a)\end{cases}$

זה קורה מכיון ש$\int\limits_{\mathbb{R}^d}\psi\cdot\delta_a dx=\int\limits_{\mathbb{R}^d}\psi d[\delta_a]=\psi(a)$ **כאשר** $\delta_a(x)=\begin{cases}1 & x=a\\ 0 & x\ne a\end{cases}$ **שיעור 5**

5. **אמידה של תוחלת שונות ואחוזון של** $\hat{F}_n$ $\begin{cases}\mu_{\hat{F}_n}=\overline{X}\\ \sigma_{\hat{F}_n}=S_n^2=\frac{1}{n}\sum\limits_{i=1}^{n}(x_i-\overline{X})^2\\ q_{\hat{F}_n}(p)=inf\{x\in\mathbb{R}\,|\,\hat{F}_n(x)>p\}=X_{(\lceil np\rceil)}\end{cases}$ **כאשר ה**t **של המשוואה** $\mathbb{P}(\hat{F}_n<t)=1-\alpha$

6. $\underset{x\in\mathbb{R}}{Sup}|\hat{F}_n-F|=||\hat{F}_n-F||_\infty\xrightarrow{\mathbb{P}/a.s}0$ משפט $Glivenko-Cantelli$

7. **אם אנו עושים מעבר מסטטיטיקה אפרמטרית לסטטיטיקה פרמטרית נשים לב כי** $\hat{F}_n(x)=F(\hat{\lambda_n},x)$**דוגמא :**
שאלה 7 תרגיל 1 אם $X_1,...X_n\overset{i.i.d}{\sim}Exp(\lambda)$ אז $E||\hat{F}_n-F||_{L_2}^2=E_F\int\limits_{\mathbb{R}}(e^{-\hat{\lambda}_nx}-e^{-\lambda x})^2dF_X$ כאשר $\hat{\lambda}$ הוא לרוב אומד $M.L.E$ או אומד בשיטה כלשהי

8. נניח שאנו רוצים להראות ש$\underset{n\to\infty}{lim}\underset{F\in\mathcal{F}}{Sup}E_F|\sqrt{n}(T(\hat{F}_n)-T(F))|^p=\infty$ עבור $p>0$
הרעיון הוא לחסום מלמטה את הביטוי $\underset{n\to\infty}{lim}\underset{F\in\mathcal{F}}{Sup}E_F|\sqrt{n}(T(\hat{F}_n)-T(F))|^p\ge\sqrt{n}\cdot a^p>0$
כאשר $a^p\in\mathbb{R}$ קבוע כלשהו נשים לב כי $E_F|(T(\hat{F}_n)-\underbrace{T(F)}_{=R})|^p\ge R^p\cdot\mathbb{P}(|T(\hat{F})-R|)$ ואז נחרטט ונגיע ל$E_F|(T(\hat{F}_n)-\underbrace{T(F)}_{=R})|^p\ge R^p$
ואז הגענו לכך ש

$$\underset{n\to\infty}{lim}\underset{F\in\mathcal{F}}{Sup}E_F|\sqrt{n}(T(\hat{F}_n)-T(F))|^p\ge\sqrt{n}\cdot a^p>0$$

אשאיף את $a^p\sqrt{n}\longrightarrow\infty$ ואקבל את הדרוש בשאלה ! **מבוחן של פבל שנה שעברה**

1.1 רווח סמך על סמך ההתפלגות האמפירית $\hat{F}_n$

1.1.1 רווח סמך אסימפטוטי על סמך פונק ההתפלגות האמפירית $\hat{F}_n(x_0)$המקרה הבדיד בנקודה קבועה x_0

לפי משפט הגבול המרכזי נקבל כי אם $X_1,...X_n\overset{i.i.d}{\sim}f$ נקבל כי

$$\sqrt{n}(\hat{F}_n(x_0)-F(x_0)|)\xrightarrow{d}N[0,F(x_0)(1-F(x_0))]\ \forall x_0\in\mathbb{R}$$

על סמך טענה זו נבנה רצועת סמך

$$\hat{F}_n(x_0)-\frac{1}{\sqrt{n}}\cdot[F(x_0)(1-F(x_0))]\cdot Z_{1-\alpha/2}<F(x_0)<\hat{F}_n(x_0)+\frac{1}{\sqrt{n}}\cdot[F(x_0)(1-F(x_0))]\cdot Z_{1-\alpha/2}\ \forall x_0$$

1.1.2 נרצה לבנות רצועת סמך עבור המקרה ש$[F(x_1),F(x_1),..F(x_n)]$ כאשר $x_1<x_2<...<x_n$ המקרה הוקטורי: $\hat{F}(\vec{x_0})$

תחילה נגדיר $(X^k)=\begin{pmatrix}(x_1)\\ \vdots\\ (x_k)\end{pmatrix},F_n(X^k)=\begin{pmatrix}\hat{F}_n(x_1)\\ \vdots\\ \hat{F}_n(x_k)\end{pmatrix},F(X^k)=\begin{pmatrix}F(x_1)\\ \vdots\\ F(x_k)\end{pmatrix}$

- מתקיים כי $\underset{F\in\mathcal{F}}{Sup}E_F||\sqrt{n}(\hat{F}(X^k)-F(X^k))||_2\le C_k$ הקבוע C תלוי בk
- לפי הCLT מתקיים כי $\sqrt{n}(\hat{F}(X^k)-F(X^k))\xrightarrow{d}N(0,\Sigma_F)$
- כאשר נגדיר את $\Sigma_F=Cov_F[\begin{pmatrix}1_{\{x_1<x\}}\\ \vdots\\ 1_{\{x_k<x\}}\end{pmatrix},\begin{pmatrix}1_{\{x_1<x\}}\\ \vdots\\ 1_{\{x_k<x\}}\end{pmatrix}]$ כאשר $[\Sigma_F]_{i.j}=E_F1_{\{X_1\le x_i\}}1_{\{X_1\le x_j\}}-E_F1_{\{X_1\le x_i\}}\cdot E_F1_{\{X_1\le x_j\}}=F(x_i\wedge x_j)-F(x_i)F(x_j)$
- נשים לב כי $||\sqrt{n}(\hat{F}(X^k)-F(X^k))||_{\Sigma_F}=\sqrt{n}(\hat{F}(X^k)-F(X^k))^T\cdot\Sigma_F^{-1/2}\cdot\Sigma_F^{-1/2}\cdot\sqrt{n}(\hat{F}(X^k)-F(X^k))\xrightarrow{d}\chi^2_{k,1-\alpha}$

נקבל מכל הטענות אללו **קבוצת סמך שהיא בצורת אלפסויד**

$$E_\alpha(X^n)=\{v\in\mathbb{R}^k\mid||\hat{F}(X^k)-v||_{\Sigma_F}\le\frac{1}{n}\chi^2_{k,1-\alpha}\}$$

כלומר $\mathbb{P}[F(X^k)\in E_\alpha]\Longrightarrow\psi[\chi^2_{k,1-\alpha}]$ הצוברת של χ^2 !

1.1.3 רווח סמך עבור המקרה הרציף של פונק ההתפלגות האמפירית לפי התפלגות קולמוגרוב ולפי משפט (DKW) כלומר עבור $F(x)$ כל הנקודות ב$\mathbb{R}$

מוטיבציה: על מנת לבנות את הרווח הסמך עבור המקרה הרציף נצטרף להעזר במספר טענות חשובות

1. אומד $\hat{F}_n$הוא **עקיב ל**F **אם מתקיים כי** $\rho(\hat{F}_n,F)\xrightarrow{p}0$ **כאשר** ρ היא מטריקה במרחב הפונקציות $\mathcal{F}$ **לרוב** $\rho=||\cdot||_\infty$
בדרך כלל הכוונה היא ש$T(\hat{F}_n)\xrightarrow{p}T(F)$ משיעורי בית של פבל !
2. **טענה חשובה :** משפט $Glivenko-Cantelli$ מתקיים כי $\underset{x\in\mathbb{R}}{Sup}|\hat{F}_n-F|=||\hat{F}_n-F||_\infty\xrightarrow{\mathbb{P}/a.s}0$

טענות חשובות על המשתנה $D_n\overset{def}{=}\sqrt{n}\cdot\underset{x\in\mathbb{R}}{Sup}|\hat{F}_n(x)-F(x)|=\sqrt{n}\cdot||\hat{F}_n(x)-F(x)||_\infty$

- D_n הוא **לא סטטיסטי** מכיון שההתפלגות שלו תלויה בF ! כאשר F רציפה ההתפגלות שלו לא תלויה בF
  - ראה תרגיל 1 כאשר F היא פונק אטומית עם $\begin{cases}\{x_1,...x_k\}\\ \{p_1,...p_k\}\end{cases}$ אז $D_n=\underset{1\le j\le k}{Max}|\hat{p}_j-p_j|$ כאשר $\hat{F}_n(x_j)=\frac{1}{n}\sum_{i=1}^n1_{\{X_i=x_j\}}=\hat{p}_j$
- **טענה חשובה:** D_n **הוא פיבוט כלומר ההתפגלות שלו לא תלויה ב**F **רק כאשר** F **רציפה !!**
- יהי $F\overset{i.i.d}{\sim}X_1,...X_n$ כאשר F רציפה $(*)$ אזי מתקיים כי $F(X_j)=V_j\sim U(0,1)$
  - **מדוע** D_n **הוא פיבוט כאשר** F **רציפה:**

$$\sqrt{n}\cdot\underset{x\in\mathbb{R}}{Sup}|\hat{F}_n(x)-F(x)|=\sqrt{n}\cdot\underset{x\in\mathbb{R}}{Sup}|\frac{1}{n}\sum_{i=1}^n1_{\{X_j\le x\}}-F(x)|=\sqrt{n}\cdot\underset{x\in\mathbb{R}}{Sup}|\frac{1}{n}\sum_{i=1}^n1_{\{F(X_j)\le F(x)\}}-F(x)|=\sqrt{n}\cdot\underset{x\in\mathbb{R}}{Sup}|\frac{1}{n}\sum_{i=1}^n1_{\{(V_j\le u\}}-u|$$

  - **אינטואציה שקרית :** $\mathbb{P}[F(X)\le x]\overset{*}{=}\mathbb{P}[X\le F^{-1}(x)]=\int_{-\infty}^{F^{-1}(x)}f_xdx=F[F^{-1}(x)]-F[-\infty]=FF^{-1}(x)=x$
  - **הוכחה פורמלית:** $\mathbb{P}[F(X)\le t]=E[1_{\{F(X)\le t\}}]\overset{*}{=}\int_{\mathbb{R}}1_{\{F(X)\le t\}}\cdot f_xdx=\int_0^11_{\{u\le t\}}\cdot du=u=F(X)$ פתרון תרגיל 1
- **נניח כי** F **רציפה אזי מתקיים כי :** $\mathbb{P}(D_n\le t)\xrightarrow{n\to\infty}\sum_{j=-\infty}^{\infty}(-1)^j\cdot e^{-2j^2t^2}$ כאשר $\sum_{j=-\infty}^{\infty}(-1)^j\cdot e^{-2j^2t^2}$ **היא התפלגות קולמוגרוב.**

רצועת סמך קולמוגקרוב אסימפטוטית :

$$\mathbb{P}[\underbrace{\hat{F}_n(x)-\frac{1}{\sqrt{n}}\cdot K_{1-\alpha}}_{L_n(x)}\le F(x)\le\underbrace{\hat{F}_n(x)+\frac{1}{\sqrt{n}}\cdot K_{1-\alpha}}_{U_n(x)}]=1-\alpha$$

כאשר $K_{1-\alpha}$ **זה אחוזון ה**$1-\alpha$ **של התפלגות קולמוגרוב**

– כלומר $D_n \xrightarrow{d} K_{1-\alpha}$ כאשר $1-\alpha$ **זה התפלגות קולמוגרוב**

- מכיון שקשה לחשב את אחוזון של התפלגות קומוגרוב נעזרים בטענה הבאה **משפט** (DKW) **מתקיים כי** $\mathbb{P}_F[\underset{x\in\mathbb{R}}{Sup}|\hat{F}-F|\geq\varepsilon]\geq 2e^{-2n\varepsilon^2}$

– בעזרתו בונים רווח סמך מבלי להתבסס עם אחוזון קומוגרוב מגדירים את $\varepsilon=\frac{c}{\sqrt{n}}$ ואז נקבל $\mathbb{P}_F[D_n\geq c]\leq 2e^{-2c^2}$ ואז

$$\mathbb{P}_F(D_n\leq c)\geq 1-2e^{-2c^2}=1-\alpha$$

* **רווח סמך על סמך משפט** DKW **:** נגדיר את $1-\alpha=1-2e^{-2c^2}$ **ונקבל את רווח הסמך**

$$\hat{F}_n(x)-\frac{1}{\sqrt{n}}\cdot c_\alpha\leq F(x)\leq\hat{F}_n(x)+\frac{1}{\sqrt{n}}\cdot c_\alpha$$

כאשר $c_\alpha=\sqrt{\frac{1}{2}log(\frac{\alpha}{2})}$ שיעור 4

- אם נשאל את השאלה האם ניתן למצוא אומד $\tilde{F}$ לF שמתכנס בקצב מהיר יותר מ$\sqrt{n}$ אז התשובה היא **שלא**

– **טענה :** $\underset{n}{liminf}\underset{F\in\mathcal{F}}{Sup}E_F||\sqrt{n}(\hat{F}_n-F)|| \overset{Claim}{=} \underset{n}{liminf}\underset{\tilde{F}\in\mathcal{F}_\mathcal{X}}{inf}\underset{F\in\mathcal{F}}{Sup}E_F||\sqrt{n}(\tilde{F}-F)||_\infty$ כלומר במונחי $mini\,max$ $\hat{F}$ הוא אומד האופטימלי!

- $D_n\overset{claim}{=}\sqrt{n}\underset{j\leq n}{Max}(|\frac{j}{n}-F(x_j)|\vee|\frac{j-1}{n}-F(x_j)|)$ **מתרגיל 1 בשיעורי בית ! נועד כדי בפועל לחשב את** D_n

הקטע הזה לא לבחינה

הוכחה של התפלגות קולמוגרוב יכולה להתבסס על משהו שנקרא תהליך מקרי $(Random\ process)$ **: ראה רישומים של פבל שבוע 1**

נגדיר תהליך סטוכסטי $x\in[0,1]$, $\overline{B}_n(x)=\sqrt{n}(\hat{F}_n(x)-x)$

1. **תאוריית דונסקר** $((Donsker))$ **:** $[\sqrt{n}(\hat{F}_n(x)-F(x)),x\in\mathbb{R}]\xrightarrow{w}[\overline{B}(F(x)),x\in\mathbb{R}]$

   (א) כאשר ההתכנסות כאן מוגדרת $E_F[\psi(\overline{B}_n)]=\underset{f\in\mathcal{G}}{\int}\psi(f)\mu_n df\longrightarrow\underset{f\in\mathcal{G}}{\int}\psi(f)\mu df$

   i. כאשר $\psi:\mathcal{G}\to\mathbb{R}$ הוא **פונקציונל רציף וחסום** ו$\mathcal{G}$ הוא מרחב כל הפונקציות.

הקטע הזה לא לבחינה

2 אמידה של פונקציונלים

הגדרה פונקציונים: $T:\mathcal{F}\to\mathbb{R}^k$ **אנו נתמקד ב**$k=1$ **ו**$\mathcal{F}$ **מרחב פונק התפלגות מצטברת.**

סוגים שונים של פונקציונלים :

1. **פונקציונל לינארי:** $T(F)=\underset{\mathbb{R}}{\int}\psi(x)dF(x)$

   (א) נשים לב כי מדובר בפונקציונל לינארית מכיון ש$T(F+G)=\underset{\mathbb{R}}{\int}\psi(x)d(F+G)(x)=\underset{\mathbb{R}}{\int}\psi(x)dF(x)+\underset{\mathbb{R}}{\int}\psi(x)dG(x)=T(F)+T(G)$

   (ב) **מקרה מעניין של פונקציונל לינארי** $\begin{cases}T(\hat{F}_n)=\underset{\mathbb{R}^d}{\int}\psi d[\hat{F}_n]=\frac{1}{n}\sum\limits_{j=1}^{n}\psi(X_j)=\overline{\psi}_n\\ T(\hat{F}_n)=\int\limits_a^b\psi d[\hat{F}_n]=\frac{1}{n}\sum\limits_{j=1}^{n}\psi(X_j)1_{\{a\leq x_j\leq b\}}\\ V_{\hat{F}_n}=\int(\psi-\overline{\psi})^2d\hat{F}_n=\int\psi^2d\hat{F}_n-(\overline{\psi})\end{cases}$

   זה קורה מכיון ש$\underset{\mathbb{R}^d}{\int}\psi\cdot\delta_a dx=\underset{\mathbb{R}^d}{\int}\psi d[\delta_a]=\psi(a)$ **כאשר** $\delta_a(x)=\begin{cases}1 & x=a\\ 0 & x\neq a\end{cases}$ **שיעור 5**

   i. נשים לב שאם $T(\hat{F}_n)=\int\limits_b^a\psi d[\hat{F}_n]=\frac{1}{n}\sum\limits_{i=1}^{n}1_{\{a\leq X_i\leq b\}}$ מסקנה מהבוחן לדוגמא של פבל

   ii. $\int\limits_b^a xd\triangle_X=x\cdot 1_{\{x\in(a,b]\}}$ **מסקנה מהבוחן לדוגמא של פבל !**

   (ג) נניח כי $||G_t-G||_\infty\xrightarrow{t\to 0}0$ וגם ψ פונקציה **חסומה** כלומר (לכל $\delta>0$ קיימת ψ_δ כך ש $||\psi-\psi_\delta||_\infty<\delta$) **אזי** מתקיים כי $\int\psi dG_t\xrightarrow{t\to 0}\int\psi dG$ מתרגול פבל

   i. **מסקנה : אם** ψ **פונק חסומה אזי** $\dot{T}(G-F)=\underset{t\to 0}{lim}\int\psi d(G_t-F)=\int\psi d(G-F)$ **מתרגול פבל**

ii. אם ψ לא חסומה **לא בהכרח קיימת נגזרת הדמר !**

iii. ראה תרגיל 2 שאלה 5 יש שם גם דוגמא נגדית מעולה ! נקח $\begin{cases} \psi(x) = x^2 \\ G_\varepsilon = \begin{cases} \varepsilon & x = \frac{1}{\varepsilon} \\ 0 & otherwise \end{cases} \end{cases}$ ואז נקבל כי $\int \psi dG_\varepsilon = \frac{1}{\varepsilon^2} \cdot \varepsilon \xrightarrow{\varepsilon \to 0} \infty$

2. **פונקציונל לא לינארי** $T(F) = h[T_1(F), T_2(F), ..., T_m(F)]$

(א) עבור $h : \mathbb{R}^m \to \mathbb{R}$ כאשר $T_j(F)$ הוא פונקציונל לינארי

(ב) דוגמא $T_2(F) - T_1(F)^2 = E[X^2] - E[X] = \sigma^2(F)$ כאשר $T_j = \int_{\mathbb{R}} x^j dF(x)$

(ג) דוגמא אם $T(F) = \int (F(x) - x)^2 dx$

אז מתקיים כי $T(\hat{F}_n) = \int (\hat{F}_n(x) - x)^2 dx = \sum_{j=1}^{n} \int_{X_{(j)}}^{X_{(j+1)}} (\frac{j}{n} - x)^2 dx = \sum_{j=1}^{n} [(\frac{j}{n})^2 (X_{(j+1)} - X_{(j)}) - (\frac{j}{n})(X_{(j+1)}^2 - X_{(j)}^2) + \frac{1}{3} X_{(j+1)}^3 - X_{(j)}^3]$

שאלה 8 תרגיל 2 אצל פבל

3. פונקציונלים לא לינאריים סמויים : מוגדרים להיות **הפתרון של המשוואה** $\int_{\mathbb{R}} h(x,t) dF(x) = 0$ **עבור** $h : \mathbb{R}^2 \to \mathbb{R}$

(א) דוגמא $q_F(p)$ עבור F רציפה מתקיים כי $\int_{\mathbb{R}} 1_{\{x \le t\}} - p dt = 0 \iff F(t) = p$

2.1 שיטת הדלתא הפרמטרית עבור מקרה חד מימדיי ועבור מקרה רב מימדיי

הערה :

נשים לב מכיון ש$ImT \subset \mathbb{R}$ נדבר על **רווחי סמך ולא על רצועות סמך**

תזכורת על רווחי סמך בשיטת הדלתא המקרה החד מימדיי והמקרה הרב מימדיי :

- **מקרה חד מימדיי :**
  - נניח כי $\hat{\theta}$אומד לθ כך ש$\sqrt{n}[\hat{\theta} - \theta] \xrightarrow{d} N(0, V(\theta))$

  כאשר $V(\theta)$**ידועה** ותהיי $g : \Theta \to \mathbb{R}$ כך שg גזירה ו$\theta \in \Theta$ אזי מתקיים כי $\sqrt{n}[g(\hat{\theta} - g(\theta))] \xrightarrow{d} N(0, V(\hat{\theta}) \cdot |g^{'}(\theta)|^2)$
  - לכן הרווח סמך יהיה

$$g(\hat{\theta}) - \frac{1}{\sqrt{n}} \cdot \sqrt{V(\hat{\theta}) \cdot |g^{'}(\hat{\theta})|^2} \cdot Z_{1-\frac{\alpha}{2}} \le g(\theta) \le g(\hat{\theta}) + \frac{1}{\sqrt{n}} \cdot \sqrt{V(\hat{\theta}) \cdot |g^{'}(\hat{\theta})|^2} \cdot Z_{1-\frac{\alpha}{2}}$$

  - אם V **לא ידועה אומדים אותה כלומר מכניסים את האומד** $\hat{\theta}$ **לשונות !**

- **מקרה רב מימדיי:**
  - נניח כי $\hat{\theta}$אומד ל $\overrightarrow{\theta}$ כך ש$\sqrt{n}[\hat{\theta} - \overrightarrow{\theta}] \xrightarrow{d} N(0, \Sigma)$ כאשר Σ**ידועה** ותהיי $g : \Theta \to \mathbb{R}$ כך שg גזירה ו$\overrightarrow{\theta} \in \Theta$ אזי מתקיים כי

$$\sqrt{n}[g(\hat{\theta}) - g(\overrightarrow{\theta})] \xrightarrow{d} N(\overrightarrow{0}, (\nabla g)^T \cdot \Sigma \cdot (\nabla g))$$

  - אם $g : \Theta \to \mathbb{R}^m$ כך ש$k = Range(\Theta) \le Range(g) = m$ אז תחת אותם תנאים בדיוק נקבל כי $\sqrt{n}[g(\hat{\theta}) - g(\overrightarrow{\theta})] \xrightarrow{d} N(\overrightarrow{0}, (Jg)^T \cdot \Sigma \cdot (Jg))$
    * כאשר $J(g)$ היא מטריצת יעקובאן $J(g) = \begin{pmatrix} | & & | \\ \nabla f_1 & .... & \nabla f_m \\ | & & | \end{pmatrix}$
    * **רווח סמך במקרה הרב מימדיי הינו**

$$g(\hat{\theta}) \pm \frac{1}{\sqrt{n}} \sqrt{\nabla g(\hat{\theta}) \cdot \sum(\hat{\theta}) \cdot [\nabla g(\hat{\theta})]^T} \cdot Z_{1-\frac{\alpha}{2}}$$

ראה דוגמא בתרגיל 2 שאלה 3

2.2 שיטת הדלתא עבור המקרה הלא פרמטרי ופונקציית רגישות:

שיטת הדלתא המקרה הלא פרמטרי !! נרצה לבצע משהו דומה למה שביצענו במקרה הפרמטרי :
טענה שיטת הדלתא במקרה הלא פרמטרי :

$$\sqrt{n}[T(\hat{F}_n)-T(F)] \overset{*}{=} \dot{T}_F(\sqrt{n}[T(\hat{F}_n)-T(F)])\cdot(1+o(1)) \overset{def}{=} \sqrt{n}\int_{\mathbb{R}}\psi d[(\hat{F}_n-F)(1+o(1))] = \sqrt{n}[\frac{1}{n}\sum_{i=1}^{n}\psi - E_F] \xrightarrow{d} N(0,V_F)$$

- נדרוש כי $\dot{T}_F(\sqrt{n}[T(\hat{F}_n)-T(F)])\cdot(1+o(1)) \overset{def}{=} \sqrt{n}\int_{\mathbb{R}}\psi d[(\hat{F}_n-F)(1+o(1))]$
- נדרוש כי $\int_{\mathbb{R}}(\psi-E_{\hat{F}_n}(\psi))^2 d[F_n] = V_{F_n} \longrightarrow V_F = \int_{\mathbb{R}}(\psi-E_F(\psi))^2 d[F]$
  - דוגמא אם $T(F)=E_f(\psi)$ אז נקבל כי $V_{\hat{F}_n} = \int_{\mathbb{R}}[\psi(x)-T(\hat{F}_n)]^2 d\hat{F}_n = \frac{1}{n}\sum_{i=1}^{n}\psi^2(x_j)-[\overline{\psi}]^2 \xrightarrow{p} V_F$
- הטענה שמוצגת כאן היא **כללית** הן עבור **פונקציונל לינארי והן עבור פונקציונל לא לינארי** עבור כל מקרה בנפרד שיטת הדלתא מתנהגת קצת שונה.
- **נקבל את הרווח סמך הבא**

$$T(\hat{F}_n) \pm \frac{1}{\sqrt{n}}\cdot Z_{1-\frac{\alpha}{2}}\sqrt{V(\hat{F}_n)}$$

כאשר $\begin{cases} V_{F_n} \longrightarrow V_F \\ V_{F_n} = \int_{\mathbb{R}}(\psi-E_{\hat{F}_n}(\psi))^2 d[F_n] \end{cases}$

  - כאשר הרבה פעמים כדי לחשב את $V(\hat{F}_n)$ נחשב את פונקציית הרגישות שהיא $\begin{cases} V(\hat{F}_n) = \int_{\mathbb{R}} L_{\hat{F}_n}(x)^2 d[\hat{F}_n] \\ L_F(x) = \dot{T}(\triangle_X - F) \end{cases}$
  - קודם נחשב את $\dot{T}(G_\varepsilon - F)$
  - לאחר מכן נציב בחישוב במקום $G_\varepsilon = \triangle_X$
  - נחשב את V_Fעם ההצבה של $G_\varepsilon = \triangle_X$
  - לבסוף נחליף את Fב$\hat{F}$בחישוב
- שימו לב שברווח סמך אני משתמש **באומד לשונות ובשיטת הדלתא הביטוי** $\sqrt{n}[T(\hat{F}_n)-T(F)] \overset{*}{=}$ **מתכנס להתפלגות נורמלית עם שונות שהיא כן ידוע** V_F

הגדרה 2.1 פונקציית רגישות : $L_F(x) = \dot{T}(\triangle_X - F)$ **כאשר** $\triangle_X = 1_{\{u\geq x\}}$
המטרה של פונקציית הרגישות היא בחישוב השונות בשיטת דלתא לפי שניתן לראות בטענה השניה !

- פונקציית רגישות עבור פונקציונל לינארי $L_F(x) = \dot{T}(\triangle_X - F) = \int_{\mathbb{R}}\psi d[\triangle_X - F] = \psi - E_F(\psi)$
  - **מסקנה עבור פונקציונלים לינארים :** $V_F = \int_{\mathbb{R}}(\psi - E_F(\psi))^2 d[F] = \int_{\mathbb{R}} L_F(x)^2 d[F]$

2.2.1 שיטת הדלתא עבור פונקציונלים לינאריים : כלומר עבור $T(F) = \int_{\mathbb{R}}\psi dF$

נגדיר נגזרת גאטו מדובר בנגזרת חלקית עבור פונקציות :

$$\dot{T}_F(F-G) = \lim_{t\to 0}\frac{T[(1-t)F+tG]-T(F)}{t} = \lim_{t\to 0}\frac{T[(F+t(G-F)]-T(F)}{t}$$

הגדרה פורמלית של נגזרת גאטו :
תהיי $T:(D,||\cdot||_\infty) \longrightarrow (E,||\cdot||_\infty)$ כאשר D מרחב של הפרש פונק התפלגות עם השתנות סופית וE מרחב פונקציות התפלגיות עם השתנות סופית.
נאמר כי T היא גזירה בנקודה $F\in D$ בכיון $G-F$ אם מתקיים כי $\dot{T}_F(F-G) = \lim_{t\to 0}\frac{T[(F+t(G-F)]-T(F)}{t}$

- במקרה של פונקציונל לינארי נקבל כי $\dot{T}_F(F-G) = \int_{\mathbb{R}}\psi d[F-G]$

ואז לפי שיטת הדלתא במקרה של **פונקציונל לינארי** נקבל כי $\sqrt{n}[T(\hat{F}_n)-T(F)]=\dot{T}_F[\hat{F}_n-F]=\int_{\mathbb{R}}\psi d[\sqrt{n}(\hat{F}_n-F)]=\sqrt{n}[\overline{\psi}-E_F(\psi)]\xrightarrow{CLT}$ $N(0,V_F)$

טענות בסיסיות על נגזרת הדמר ופונקציונלים לינאריים:

1. נניח כי $||G_t-G||_\infty\xrightarrow{t\to 0}0$ וגם ψ פונקציה **חסומה** כלומר (לכל $\delta>0$ קיימת ψ_δ כך ש $||\psi-\psi_\delta||_\infty<\delta$) **אזי** מתקיים כי $\int\psi dG_t\xrightarrow{t\to 0}\int\psi dG$ מתרגול פבל

   (א) **מסקנה : אם ψ פונק חסומה אזי** $\dot{T}(G-F)=\lim_{t\to 0}\int\psi d(G_t-F)=\int\psi d(G-F)$ **מתרגול פבל ! ומתרגיל 2 שאלה 5 !**

   (ב) **אם ψ לא חסומה לא בהכרח קיימת נגזרת הדמר !** ראה דוגמא נגדית מתרגול פבל מספיק לבדוק עבור $\begin{cases}\psi(x)=x\\ G_t=(1-t)G+t\triangle_X\end{cases}$

   (ג) **דוגמא לפונק שהיא לא גזירה הדמר אבל גזירה גאטו :** ראה תרגיל 2 שאלה 5 יש שם גם דוגמא נגדית מעולה ! נקח $\begin{cases}\psi(x)=x^2\\ G_\varepsilon=\begin{cases}\varepsilon & x=\frac{1}{\varepsilon}\\ 0 & otherwise\end{cases}\end{cases}$

   ואז נקבל כי $\int\psi dG_\varepsilon=(\frac{1}{\varepsilon})^2\cdot\varepsilon\xrightarrow{\varepsilon\to 0}\infty$

2.2.2 שיטת דלתא עבור פונקציונלים לא לינארים:

כאן נצטרך להגדיר נגזרת הדמר שהיא דומה לגאטו מאוד :

תהיי $T:(D,||\cdot||_\infty)\longrightarrow(E,||\cdot||_\infty)$

- כאשר D מרחב של **הפרש פונק התפלגות** עם השתנות סופית וE מרחב פונקציות התפלגיות עם השתנות סופית.
- **המטרה של נגזרת הדמר היא לחשב את פונקציית הרגישות בקלות ואותה צריך בשביל לחשב את השונות לשיטת הדלתא ! מסקנות מהשיעורי בית !**
  - זו הסיבה שדורשים $||G_\varepsilon-G||_\infty\xrightarrow{t\to 0}0$ שהרי $||\hat{F}_n-F||_\infty\xrightarrow{n\to\infty}0$! ונצטרך זאת כאשר נשתמש בשיטת הדלתא ! מסקנות משיעורי בית

נאמר כי T היא גזירה אדמר בנקודה $F\in D$ בכיוון G_t לכל G_t שמקיים כי $||G_\varepsilon-G||_\infty\longrightarrow 0$ **אם מתקיים כי**

$$\dot{T}_F(F-G_\varepsilon)=\lim_{\varepsilon\to 0}\frac{T[(1-\varepsilon)F+\varepsilon G_\varepsilon]-T(F)}{\varepsilon}=\lim_{\varepsilon\to 0}\frac{T[(F+\varepsilon(G_\varepsilon-F)]-T(F)}{\varepsilon}$$

נשים לב שזה שקול ל

$$||\frac{T[(F+t(G_\varepsilon-F)]-T(F)}{\varepsilon}-\dot{T}(F-G)||_\infty\longrightarrow 0$$

- **חובה לבדוק בנגזרת הדמר כי** $||G_t-G||_\infty\longrightarrow 0$ **!! זה ממש קריטי !**
  - **ראה דוגמא בבוחן של פבל !**
- אז לפי לשיטת הדגלתא במקרה של פונקציונל לא לינארי נקבל כי : $\sqrt{n}[T(\hat{F}_n)-T(F)]\xrightarrow{d}\dot{T}_F[\overline{B}\circ F]\sim N(0,V_F)$
  - **כאשר גם כאן** $V_F=\int_{\mathbb{R}}L_F(x)^2d[F]$.
- **נגזרת הדמר של אחוזון כלומר של הפונקציונל** $F^{-1}(q)=q_p(F)=Min\{x\in\mathbb{R}\,|\,F(x)\geq p\}$ **הינה** $\dot{T}(G-F)=\lim_{t\to\infty}\frac{q_t-q}{t}=\frac{F(q)-G(q)}{f(q)}$
  - באחוזון הגדרנו את האחוזון q_t להיות $(1-t)F(q_t)+tG_t(q_t)=p$ כאשר נדרוש כי $\begin{cases}G_t\ continuous\\ ||G_t-G||_\infty\\ F\ continuous\end{cases}$
  - מתקיים כי $q_t\longrightarrow q$ מוכיחים טענה זו !
  - **עבור אחוזון האומד שנקבל הינו :** $T(\hat{F}_n)=min\{x\in\mathbb{R}\,|\,\sum_{i=1}^{n}1_{\{X_i<x\}}>pn\}=X_{(\lceil np\rceil)}$
  - **רווח סמך של אחוזון תחת הדרישות לעיל הינו**

$$X_{(\lceil np\rceil)}\pm\frac{1}{\sqrt{n}}\sqrt{p(1-p)}\cdot\frac{1}{\hat{f}_n(X_{(\lceil np\rceil)})}\cdot Z_{1-\frac{\alpha}{2}}$$

כאשר $V_F=\cdot\frac{p(1-p)}{[\hat{f}_n(X_{(\lceil np\rceil)})]^2}=\frac{F(q)[1-F(q)]}{[F(q)]^2}$**מתרגול פבל.**

2.2.3 פונק רגישות לפונקציונלים שונים :

נחשב את פונק הרגישות ב3 שלבים :

- נחשב את $\underset{t\to0}{lim}\frac{T[(F+t(G_t-F)]-T(F)}{t} = \underset{t\to0}{lim}\frac{T[(1-t)F+tG_t]-T(F)}{t} = \dot{T}_F(F-G_t)$ עבור F ממש !!
- לאחר מכן מצהיר כי $G_t=\delta_x$
- נציב במקום F את $\hat{F}$אם נרצה לחשב את $V_{\hat{F}_n}$ בשביל רווח סמך אסימפטוטי

1. פונקציית רגישות עבור **פונקציונל לינארי** מהצורה $T(F)=\int xdF$ נקבל כי $L_F(x)=\dot{T}(\triangle_X-F)=\int\limits_{\mathbb{R}}\psi d[\triangle_X-F]=\psi-E_F(\psi)$
2. פונקציית רגישות עבור **אחוזון** $q=T(F)=min\{x\in\mathbb{R}\,|\,F(x)>p\}=F^{-1}(p)$ מהצורה $T(F)=\int xdF$ נקבל כי $L_F(x)=\dot{T}(\triangle_X-F)=\frac{p-1_{\{q\geq x\}}}{f(q)}$

סיוע בחישוב פונק רגישות $L_F(x)$: **ראה דוגמא בפתרון בוחן לדוגמא משנה שעברה !**
במידה ויש לנו פונקציונל מסובך ניתן להציג אותו בעזרת 2 פונקציונלים פשוטים שלהם אנו יודעים לחשב את $L_F(x)$
במקרה זה נגדיר $h\circ g(x)=h(T_0,T_1)=T_{complex}$ כאשר $g(x,y)=(T_0,T_1)$ ונקבל כי

$$\dot{h}(G_\varepsilon-F)=D(h\circ g)=Dh(g)\circ D(g)<\begin{pmatrix}\frac{dh}{dT_0}\\ \frac{dh}{dT_1}\end{pmatrix},\begin{pmatrix}L_{T_0(F)}\\ L_{T_0(F)}\end{pmatrix}>=\frac{dh}{dT_0}\cdot L_{T_0(F)}+\frac{dh}{dT_1}\cdot L_{T_1(F)}$$

3 שיטת $Bootstrap$ לאמידה של פונקציונלים ורווחי סמך.

מוטיבציה והערה חשובה :

- הרעיון הוא כזה נניח ביצענו איזושהי אמידה לבעיה כלשהי וקיבלנו אומד נרצה **לבחון אותו ואולי לשפר אותו על ידי הפחת ה**$bias$ **או הפחתת השונות.**
- נרצה להשתמש בבוטסארפ אם נרצה **לאמוד את השונות של האומד ולא נוכל לעשות זאת בצורה סגורה.**

שיטת ה$Bootstrap$
נניח שישנו $T(F)$ פונקציונצל הצבה ואנו מעוניינים לאמוד את ההטייה שלו כלומר את $Bias=E[T(F)]-T(F)$ או כל פונקציה אחרת למשל Ψ
איך לקבל אומדי $Bootstrap$ הרבה פעמים נשתמש בשיטת ה$Bootstrap$ כדי לאמוד את השונות או את ה$bias$ של אומדי $M.L.E$
או הטייה של אומד לשונות וכולי
ובשלב הII**:** נחסיר מהאומד $T(\hat{F}_n)$ את הטייה **ונקבל אומד** $Bootstrap$עם הטייה נמוכה יותר לפעמים כלומר $T(\hat{F}_n)-\hat{t}_n\overset{def\ Bootestimator}{=}\hat{T}_n^B$
דוגמא הטייה של פונקציונל $T(F)$ הינה $t=E_FT(F)-T(F)$
נניח ויצא לנו כי $\hat{t}_n=E_{\hat{F}_n}T(F_n^*)-T(F_n^*)$ואז **אומד** $Bootsrap$ **עם הטייה מופחת הינו** $T(F)-\hat{t}_n$
היא מופחת מאשר ההטייה הטבעית של אותו האומד ראה דוגמא ברישומים של פבל או ראה דוגמא שאלה 4 בשיעורי בית !

1. נעזר באומד ההצבה $\hat{F}_n$ ונגדיר $b_T(F)=ET(\hat{F})-T(F)$ או באופן כללי $E_F\Psi(F,\hat{F}_n,t)=0$ כאשר כאן $t=b_T(F)$
2. אחרי שהגדרנו את המשוואה שלנו נחליף אותה בצורה הבאה $E_F\Psi(F,\hat{F}_n,t)=0\Longleftarrow E_{\hat{F}_n}\Psi(\hat{F}_n,F_n^*,t)=0$
   (א) כעת מבחינתנו $\hat{F}_n$ ידועה וF_n^* לא ידועה כלומר $X_1^*,X_2^*,...X_n^*|(\hat{F}_n)\overset{i.i.d}{\sim}F_n^*$
      i. שימו לב כי $X_1^*,...X_n^*$ מתפלג אחיד $\frac{1}{n}$ על ערכי המדגם המקורי $X_1,...,X_n$!
   (ב) ניתן לדייק את $E_{\hat{F}_n}\Psi(\hat{F}_n,F_n^*,t)=0$ בעזרת שיטת מונטה קארלו
      i. נגריל $F_n^{*(1)}=X_1^*,X_2^*,...X_n^*|(\hat{F}_n)\overset{i.i.d}{\sim}F_n^*$ M פעמים ונקבל $F_n^{*(1)},F_n^{*(2)},...,F_n^{*(M)}$
      ii. לפי החוק החזק של המספרים הגדולים נקבל כי $\frac{1}{M}\cdot\sum\limits_{i=1}^{M}\Psi(\hat{F}_n,F_n^{*(j)},t)\overset{M\to\infty}{\longrightarrow}E_{\hat{F}_n}\Psi(\hat{F}_n,F_n^*,t)$
3. המטרה הסופית בשיטת ה$Bootstrap$ זה להגיע מ$T(F_n^*)$ל**ביטוי שתלוי רק במדגם המקורי** $X_1,..X_n$ **!**

- **דוגמא פרקטית : נפעיל את שיטת ה**$bootstrap$ **כדי לאמוד את ה**$bias$ **של אומד ההצבה**
  נניח נתון אומד $T(F)=[\int\limits_{\mathbb{R}}xdF]^2$ ונתון אומד הצבה $T(\hat{F}_n)=[\int\limits_{\mathbb{R}}xd\hat{F}_n]^2=[\overline{X_n}]^2$
  ואנו רוצים לאמוד את ההטייה של אומד ההצבה.
- דוגמא נוספת זה שימוש בשיטת ה$Bootstrap$כדי לאמוד את השונות של הפונקציונל.

למה שיטת ה$Bootstrap$**עובדת : שימוש ב2 המשפטים הבאים :**

- נניח כי Tגזירה הדמר עם פונק רגישות $L_F(x)$המקיימת כי $V_F = \int\limits_{\mathbb{R}} L_F(x)^2 dF < \infty$ וF פונק רציפה אזי

$$\underset{u\in\mathbb{R}}{Sup}|\mathbb{P}_F[\sqrt{n}[T(\hat{F}_n)-T(F)]<u]-\mathbb{P}_{\hat{F}_n}[\sqrt{n}[T(F_n^*)-T(\hat{F}_n)]<u]| \xrightarrow{\mathbb{P}} 0$$

- **א"ש ברי אנסן עידון של CLT:** $\underset{x\in\mathbb{R}}{Sup}|\mathbb{P}_F(\sqrt{n}\cdot\frac{\overline{X}-\mu}{\sigma_F}\le x)-\phi(x)|\le\frac{1}{\sqrt{n}}\cdot\frac{\rho}{\sigma_F^3}\cdot C$ כאשר $\begin{cases} C\in\mathbb{R} \\ \rho=E_F|X_1-\mu_F|^3<\infty\end{cases}$ מספר קבוע !

- ראה הרצאה של פבל

אומד $Bootstrap$ לשונות ולהטייה :

- אומד $Bootstrap$להטייה : $\hat{t}_n = E_{\hat{F}_n}[T(\hat{F}_n^*)]-T(\hat{F}_n)$ עבור פונקציונל T כלשהו
- אומד $Bootstrap$לשונות : $\hat{t}_n = E_{\hat{F}_n}[T(\hat{F}_n^*)]^2-[E_{\hat{F}_n}T(\hat{F}_n)]^2$ עבור פונקציונל T כלשהו

3.1 שיטת ה$Bootstrap$ לבנייה של רווחי סמך :

- **שיטה I:** רווח הסמך הינו $T(\hat{F}_n)\pm\hat{t}_n$ שזה שקול למשוואה

$$\mathbb{P}_F[|T(\hat{F}_n)-T(F)|<t]=1-\alpha \iff \begin{cases}E_F\psi(\hat{F}_n,F,t)=0\\ \psi=1_{\{|T(\hat{F}_n)-T(F)|<t\}}\end{cases}$$

כאשר $\hat{t}_n$ הוא פתרון המשוואה $E_{\hat{F}_n}\psi(F_n^*,\hat{F}_n,t)=0$

דוגמא ישומית נחשב רווח סמך ל$X_1,..X_n\sim N(\mu,\sigma^2)$ עבור הפונקציונל $T(F)=\int xdF$ בעזרת אומדי $M.L.E$ שהם $\hat{\theta}=\begin{pmatrix}\overline{X}\\ S_n^2\end{pmatrix}$:

1. נפתור את המשוואה $E_{\hat{F}_n}\psi(F_n^*,\hat{F}_n,u)=0$ שהיא שקולה ל

$$\mathbb{P}_F[\frac{\sqrt{n}}{S_n}\cdot|T(F_n^*)-T(\hat{F})|<\sqrt{n}\cdot\frac{t}{S_n}]=1-\alpha \iff \mathbb{P}_F[|T(F_n^*)-T(\hat{F})|<u]=1-\alpha$$

כלומר $\sqrt{n}\cdot\frac{t}{S_n}=Z_{1-\frac{\alpha}{2}} \iff \hat{t}\overset{def}{=}\frac{Z_{1-\frac{\alpha}{2}}}{\sqrt{n}\cdot S_n}$

2. לכן הרווח סמך שלנו הינו

$$T(\hat{F}_n)\pm\frac{Z_{1-\frac{\alpha}{2}}}{\sqrt{n}\cdot S_n}$$

(א) **הערה בלבד :**
נציב את התוצאה במשוואה המקורית $E_F\psi(\hat{F}_n,F,\hat{t})=0$ ונקבל $\mathbb{P}_F[\sqrt{n}\cdot\frac{|T(\hat{F}_n)-T(F)|}{S_n}<Z_{1-\frac{\alpha}{2}}]=1-\alpha$ שזה $G_{n-1,1-\frac{\alpha}{2}}[Z_{1-\frac{\alpha}{2}}]$ כאשר $G_{n-1,1-\frac{\alpha}{2}}$ היא התפלגות t עם $n-1$ דרגות חופש.

i. ואז נקבל את רווח הסמך הבא : $T(\hat{F}_n)\pm\frac{G_{n-1,1-\frac{\alpha}{2}}[Z_{1-\frac{\alpha}{2}}]}{\sqrt{n}\cdot S_n}$

- **שיטה II : נקראת** $studentized\ bootstrap\ interval$
נביט במשוואה $\mathbb{P}_F[\frac{|\mu_{F_n}-\mu_F|}{\sigma_F}\le t]=1-\alpha$ נמיר אותה ל$\mathbb{P}_{F_n}[\sqrt{n}\cdot\frac{|\mu_{F_n^*}-\mu_{\hat{F}_n}|}{\sigma_{\hat{F}_n}}\le\sqrt{n}\cdot t]=1-\alpha$ ונקבל כי $\hat{t}_n=\frac{1}{\sqrt{n}}\cdot G_{n-1}(1-\alpha)$

דוגמאות ישומיות :

- **רווח סמך בשיטת ה$Bootstrap$ עבור** $T(F)=F(x_0)$ במקרה זה נקבל $\iff \mathbb{P}_{\hat{F}_n}(n\hat{F}_n-nt\le\sum_{i=1}^{n}1_{\{X_i^*\le x\}}\le n\hat{F}_n+nt)=1-\alpha$
$\mathbb{P}_{\hat{F}_n}(|\hat{F}_n^*(x_0)-\hat{F}_n(x_0)|\le t)=1-\alpha$
ואז נקבל $\sum\limits_{\substack{i=1\\ K:|\frac{k}{n}-\hat{F}(x_0)|}}^{\infty}\begin{pmatrix}n\\k\end{pmatrix}\hat{F}(x_0)^k(1-\hat{F}(x_0))^{n-k}$ ראה שאלה 7 ברישומים של פבל

- **רווח סמך בשיטת ה$Bootstrap$ עבור** $T(F)=\int xdF$ ו$F\sim exp(\lambda)$ נקבל כי $\mathbb{P}_F[\sqrt{n}\cdot|\frac{\overline{X_n}-\mu_F}{\overline{X_n}}|\le t]=1-\alpha \Leftarrow \mathbb{P}_F[\sqrt{n}\cdot|\frac{\overline{X_n^*}-\overline{X}}{\overline{X_n^*}}|\le t]=1-\alpha$
ונקבל כי $v_n[y\in\mathbb{R}^+\,|\,|1-\frac{n}{y}|\le t]=1-\alpha$ מתפלג גמא $(n,1)$ ראה תרגיל 3 שאלה 8 אצל פבל ! והרווח סמך הינו $T(\hat{F}_n)\pm\hat{t}_n$

3.2 סיוע בחישובים בשיטת $Bootstrap$:

- $X_1^*, X_2^*, ...X_n^* \overset{i.i.d}{\sim} U[X_1, ...X_n]$ עם הסתברות $\frac{1}{n}$
- $E[S_n^2] = E[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2] = (1-\frac{1}{n})\sigma^2$
- $E[\frac{n}{n-1}\cdot S_n^2] = E[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2] = \sigma^2$

1. **חישוב בסיסי שלי תוחלת ושונות של** X_1^*:
$$\begin{cases} E_{\hat{F}_n}(X_1^* - X_1)^2 = Var(X_1^*) = S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - (\overline{X})^2 \\ E_{\hat{F}_n}(X_1^*) = \overline{X_n} \end{cases}$$

2. **חישוב תוחלת ושונות של** $\overline{X_1^*}$:
$$\begin{cases} E_{\hat{F}_n}(\overline{X_1^*}) = E_{\hat{F}_n}(X_1^*) = \overline{X_n} \\ Var(\overline{X_1^*}) = \frac{Var(X_1^*)}{n} = \frac{1}{n}\cdot S_n^2 = \frac{1}{n^2}\sum_{i=1}^{n}(X_i-\overline{X})^2 \end{cases}$$

3. **חישוב מומנטים של** X_1^*: $E_{\hat{F}_n}[\frac{1}{n}\sum_{i=1}^{n}(X_i^*)^k] \overset{i.i.d\, samples}{=} E_{\hat{F}_n}(X_1^*)^k = \frac{1}{n}\sum_{i=1}^{n}x_i^k = \overline{X^k}$
   (א) עבור $k=2$ נקבל $E_{\hat{F}_n}(X_1^*)^2 = S_n^2 + \overline{X_n}^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 = \overline{X^2}$

4. **חישוב מומנטים של ממוצע** $(\overline{X_1^*})^k$:

$$E_{\hat{F}_n}(\overline{X_1^*}^k) = E_{\hat{F}_n}(\overline{X_1^*} - \overline{X_n} + \overline{X_n})^k = E_{\hat{F}_n}[\binom{n}{0}(\overline{X_1^*}-\overline{X_n})^n + \binom{n}{1}(\overline{X_1^*}-\overline{X_n})^{n-1}\cdot\overline{X_n} + \binom{n}{2}(\overline{X_1^*}-\overline{X_n})^{n-2}\cdot(\overline{X_n})^2 + ... + (\overline{X_n})^n]$$

   (א) בשלב הבא נחפש ביטויים מהצורה $E_{\hat{F}_n}(\overline{X_n^*} - \overline{X_n})\cdot\overline{X_n}$ שהרי מתקיים כי $E_{\hat{F}_n}(\overline{X_n^*} - \overline{X_n})\cdot\overline{X_n} = 0$
   (ב) לרוב k לא יהיה גדול כל כך ונביט ב$E_{\hat{F}_n}[\overline{X_1^*} - \overline{X}]^3 = E_{\hat{F}_n}[\frac{1}{n}\sum_{i=1}^{n}(x_i^* - \overline{X})]^3 = \frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(x_i^* - \overline{X})\cdot(x_j^* - \overline{X})\cdot(x_k^* - \overline{X})$
   (ג) נביט במקרים $\begin{cases} i=j=k \\ i=j, i\neq k \\ i\neq j\neq k \end{cases}$ ונחפש ביטוים כלשהם כפול $E_{\hat{F}_n}(\overline{X_n^*} - \overline{X_n})\cdot$ והם התאפסו
   (ד) **חישובים סבוכים במומנטים** $k=3, k=4$:
      i. $E_{\hat{F}_n}(\overline{X_n^*} - \overline{X_n})^3 = \frac{1}{n^2}\cdot[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X_n})^3]$
      ii. $E_{\hat{F}_n}(\overline{X_n^*} - \overline{X_n})^4 = \frac{1}{n^3}\cdot[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X_n})^4] + \frac{3n(n-1)}{n^4}\cdot[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X_n})^2]^2$
   (ה) **עבור** $k=2$ **נקבל** $E_{\hat{F}_n}(\overline{X_n^*}^2) = E_{\hat{F}_n}[\frac{1}{n}\sum_{i=1}^{n}x_i^*]^2 = Var_{\hat{F}_n}(\overline{X^*}) + [E_{\hat{F}_n}(\overline{X^*})]^2 = \frac{1}{n}S_n^2 + (\overline{X_n})^2$

3.3 שיטת $Jake-knife$ להקטנת ההטייה מתרגיל 3 שאלה 5

יהי T אומד ל$\theta(F)$ כך ש$E_FT - \theta = \frac{a_F}{n} - \frac{b_F}{n^2} + O(n^{-3})$
כאשר a_F, b_F הן פונקציות של המדגם
נגדיר אומד להטייה בשיטת $Jake-knife$

$$\hat{B}_n^J = (n-1)(\overline{T}_n - T_n)$$

כאשר $\overline{T_n} = \frac{1}{n}\sum_{i=1}^{n}T_{n/j}$
אזי מתקיים כי

1. $E_F\hat{B}_n^J = \frac{a_F}{n} + O(n^{-2})$
2. אם נגדיר $T_n^J = T_n - \hat{B}_n^J$ שזהו אומד $Jake-knife$ נקבל כי $E_FT_n^J - \theta = O(n^{-2})$

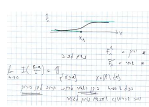
דוגמא ישומית של שיטת $Jake-knife$:
נניח ואנו רוצים לאמוד את $\theta=\mu^2$ אז אם $T=\overline{X^2}$ נקבל מסעיף א כי $E_FT-\theta=\frac{Var(x_1)}{n}-\frac{0}{n^2}+O(n^{-3})$
ואומד $Jake-knife$ יהיה

$$\overline{X^2}-\hat{B}_n^J=\overline{X^2}-(n-1)(\overline{T}_n-T_n)=\overline{X^2}-(n-1)(\underbrace{\frac{1}{n}\sum_{i=1}^{n}(\frac{1}{n-1}\sum_{i\neq j}^{n}x_i)^2}_{\overline{T_n}}-\overline{X^2})$$

4 אומדי $Kernel$ לפונקציית הצפיפות

נרצה לפתור את הבעיה הבאה בסטטיטיקה $p\in\mathbb{R}\overset{i.i.d}{\sim}X_1,...X_n$ כאשר p היא פונק צפיפות לא יודועה כלומר מניחים כי $\frac{d}{dx}F_x=p$
נשים לב כי אם באופן נהיבי נבחר לאמוד את F_x בעזרת $\hat{F}_n(x)=\frac{1}{n}\sum\limits_{i=1}^{n}1_{\{X_i\leq x\}}$ נגלה כי אין משמעות לנגזרת ומדובר בביטוי חסר משמעות
נשתמש ברעיון דומה :
נגדיר $\tilde{F}_{n,h}(x)=\frac{1}{n}\sum\limits_{i=1}^{n}I(\frac{x-X_j}{h})$ היא נקראת $``smoothed"\ empirical\ distribution\ function$
כאשר $I:\mathbb{R}\longrightarrow[0,1]$ עם תכונות של פונק התפלגות והיא גזירה ותתנהג כמעט כמו האינדיקטור !
כלומר $II(\frac{x-X_j}{h})$ היא **פונק שמקיימת את התכונות הבאות:**

- I גזירה
- I פונק עולה חלש
- $\begin{cases}\lim\limits_{x\to-\infty}I(x)=0\\ \lim\limits_{x\to\infty}I(x)=1\end{cases}$
- **נגדיר h להיות רוחב פס שמקיים כי** $\lim\limits_{h\to0}I(\frac{x-X_j}{h})=1_{\{X_j\leq x\}}$


איור 1 :

-

טענה מרכזית : מתקיים כי $\tilde{F}_{n,h}(x)\overset{h\to0}{\longrightarrow}\hat{F}_n(x)$ **והיתרון כאן ברור הוא ש**$\tilde{F}_{n,h}(x)$ **גזירה !**
נגדיר אומד חדש לp שהוא נקרא $Kernel\ Density\ Estimatior$

$$\hat{p}(x)=\frac{1}{hn}\sum_{i=1}^{n}k(\frac{x-X_j}{h})$$

כאשר k היא פונק קרנל שמקיימת

- $K(x)=\frac{d}{dx}I(x)$
- $\int\limits_{\mathbb{R}}K(x)=1$
- ניתן לדרוש תכונות נוספות מK שהרי הוא פונקציה לבחירתנו
- **נאמר כי K היא מסדר lאם מתקיים כי**
  - $\begin{aligned}&\int\limits_{\mathbb{R}}u^iK(u)du=0,\ i=1,...,l\\ &\int\limits_{\mathbb{R}}|K(u)|\cdot|u|^{l+1}du<\infty\end{aligned}$
  - אנו צריכים את ההגדרה הזו כדי לחסום את ההטייה במדד $MSE(\hat{p}(x_0),p(x_0))$

דוגמאות לפונק קרנלים נפוצות:

1. $K(u)=\frac{1}{2}1_{\{|u|\leq1\}}$ אחידה