סטטיסטיקה 2

מטלה 2

עידן פוגרבינסקי – 325069565

עילי אבני 212778229

שאלה 3

על פי ההגדרה $p-value=\inf\{\alpha: T(x^n)\in R_\alpha\}$, ניזכר בכך שאזור דחייה הוא מקיים $R_\alpha=(C_n,\infty)$

כך מתקיים $p-value=\inf\{\mathrm{a}: \mathrm{T}(\mathrm{x}^{\mathrm{n}})\geq C_n\}$

בשילוב עם הטענה הקודמת עם ההגדרה לגודל המבחן, נראה כי גודל המבחן הוא $\sup\limits_{\theta\in\Theta_0} P_\theta\big(T(X^n)\geq C_p\big)$,

כלומר הסופרימום על ההסתברות שהמבחן חורג לאזור הדחייה תחת משפחת $\theta\in\Theta_0$.

כעת ניזכר כי אנו דוחים את השארת האפס אממ $T(x^n)\geq C_p$, **אזי** $\boldsymbol{\sup\limits_{\theta\in\Theta_0} P_\theta\big(T(X^n)\geq T(x^n)\big)\leq p}$

כאשר p גודל המבחן.

נניח בשלילה כי $\sup\limits_{\theta\in\Theta_0} P_\theta\big(T(X^n)\geq T(x^n)\big)\neq p$ מאחר שהראנו את הטענה הקודמת, יתכן רק

ש$p<\sup\limits_{\theta\in\Theta_0} P_\theta\big(T(X^n)\geq T(x^n)\big)$.

נגדיר מבחן בו דוחים אממ $T(X^n)\geq T(x^n)$, מאחר ש$p<\sup\limits_{\theta\in\Theta_0} P_\theta\big(T(X^n)\geq T(x^n)\big)$, מצאנו מבחן בו

נדחה את H_0 ברמה $\sup\limits_{\theta\in\Theta_0} P_\theta\big(T(X^n)\geq T(x^n)\big)$ קטנה מ p, בסתירה לכך שרמת המבחן היא p

ולכן $\sup\limits_{\theta\in\Theta_0} P_\theta\big(T(X^n)\geq T(x^n)\big)=p=p-value$

שאלה 5

נתחיל לרשום את הביטויים ולפתח אותם, עד שנגיע לתוצאה המבוקשת

$$\beta(\theta_*) = P\left(|W| > z_{\frac{\alpha}{2}} \middle| \theta_*\right) = P\left(\left|\frac{\hat{\theta}-\theta_o}{\widehat{se}}\right| > Z_{\frac{\alpha}{2}} | \theta_*\right) = P\left(\left|\frac{\hat{\theta}-\theta_*+\theta_*-\theta_0}{\widehat{se}}\right| > Z_{\frac{\alpha}{2}} | \theta_*\right) =$$

$$= P\left(\frac{\hat{\theta}-\theta_*}{\widehat{se}} + \frac{\theta_*-\theta_0}{\widehat{se}} > Z_{\frac{\alpha}{2}}\right) + P\left(\frac{\hat{\theta}-\theta}{\widehat{se}} + \frac{\theta_*-\theta_0}{\widehat{se}} < -Z_{\frac{\alpha}{2}}\right) =$$

$$= P\left(\frac{\hat{\theta}-\theta_*}{\widehat{se}} > Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) + P\left(\frac{\hat{\theta}-\theta}{\widehat{se}} < -Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right)$$

$$= 1 - \Phi\left(Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) + \Phi\left(-Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right)$$

נימוק למעבר השלישי $\frac{\hat{\theta}-\theta_*}{\widehat{se}}$ מתפלג נורמלי

סעיף ב

נציין כי תחת הנחת האפס מתקיים: $\hat{\theta} \xrightarrow[n\to\infty]{} \theta_*$ וגם $\frac{\hat{\theta}-\theta_0}{\widehat{se}}$ אסימפטוטית נורמאלי בשילוב הטענות הללו מתקיים $\widehat{se} \xrightarrow[n\to\infty]{} 0$

ולכן $\left|\frac{\theta_0-\theta_*}{\widehat{se}}\right| \xrightarrow[n\to\infty]{} \pm\infty$ כאשר $\theta_0 \neq \theta_*$.

נפצל למקרים:

$$\theta_0 < \theta_*$$

$$\frac{\theta_0-\theta_*}{\widehat{se}} \xrightarrow[n\to\infty]{} -\infty \Longrightarrow \lim_{n\to\infty} \Phi\left(Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) = \underset{n\to\infty}{\Phi(-\infty)} = \lim \Phi\left(-Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) = 0$$

$$\lim_{n\to\infty} \beta(\theta_*) = \lim_{n\to\infty} 1 - \Phi\left(Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) + \Phi\left(-Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) = 1 - 0 + 0 = 1$$

$$\theta_0 > \theta_*$$

- $\frac{\theta_0-\theta_*}{\widehat{se}} \xrightarrow[n\to\infty]{} \infty \Longrightarrow \lim_{n\to\infty} \Phi\left(Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) = \Phi(\infty) = \lim_{n\to\infty} \Phi\left(-Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) = 1$

$$\lim_{n\to\infty} \beta(\theta_*) = \lim_{n\to\infty} 1 - \Phi\left(Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) + \Phi\left(-Z_{\frac{\alpha}{2}} + \frac{\theta_0-\theta_*}{\widehat{se}}\right) = 1 - 1 + 1 = 1$$

ולכן $\lim_{n\to\infty} \beta(\theta_*) = 1$ כאשר $\theta_0 \neq \theta_*$.

נתחיל ממציאת אנ"מ.

$$\mathcal{L}(x,p)=\prod_{i=0}^{n} f(x_i) = \prod (1-p)^{x_i} p$$

$$\log(\mathcal{L}(x,p)) = \sum_{i=0}^{n}\log\left((1-p)^{x_i-1}\right) + \sum_{i=0}^{n}\log(p) = \log(1-p)\left(\sum_{i=1}^{n} x_i - n\right) + n\log(p)$$

$$\frac{\partial \log(\mathcal{L}(x,p))}{\partial p} = \frac{-1}{1-p}\cdot\sum_{i=1}^{n} x_i + \frac{n}{1-p} + \frac{n}{p} = 0$$

$$n - np = p\sum_{i=1}^{n} x_i - np \iff \frac{n}{p} = \frac{1}{1-p}\cdot\left(\sum_{i=1}^{n} x_i - n\right)$$

$$\Downarrow$$

$$\frac{n}{\sum_{i=1}^{n} x_i} = \hat{p} \Leftarrow$$ אנ"מ

מאחר שהאנ"מ מתפלג אסימפטוטית נורמלית

ניתן לחשב את רווח הסמך של: $\hat{p} \pm z_{\frac{\alpha}{2}}\cdot \hat{se}$.

נחשב את $\hat{se}$ באמצעות

$$\hat{se} = \sqrt{\frac{1}{I_n(p)}}$$

נחפש את $I_n(p)$ ובשביל זה נמצא קודם

את $E_x[S'(p,x)]$ בעמוד הבא

כמו שבגזירה
ממקודם

$$S(x,p)=\frac{d}{dp}\log\left((1-p)^{x}p\right)=\frac{1}{p}+\frac{1-x}{1-p}$$

$$S'(x,p)=-\frac{1}{p^2}+\frac{1-x}{(1-p)^2}\Rightarrow E_x\left[-S'(x,p)\right]=E_x\left[\frac{1}{p^2}-\frac{1-x}{(1-p)^2}\right]$$

לינאריות של
תוחלת

$$=\frac{1}{p^2}-\frac{1-E_x[x]}{(1-p)^2}=\frac{1}{p^2}-\frac{1-\frac{1}{p}}{(1-p)^2}=$$

$$\Downarrow$$

$$I_n(p)=n\cdot I(\hat{p})=n\cdot\frac{1}{\hat{p}^2}-n\frac{1-\frac{1}{\hat{p}}}{(1-\hat{p})^2}=n\frac{1}{\frac{n^2}{\left(\sum_{i=1}^{n}x_i\right)^2}}-n\frac{1-\frac{\sum x_i}{n}}{\left(1-\frac{n}{\sum x_i}\right)^2}$$

$$=\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}-n\frac{1-\frac{\sum x_i}{n}}{\left(1-\frac{n}{\sum x_i}\right)^2}$$

$$\Downarrow$$

$$\hat{p}\pm Z_{\frac{\alpha}{2}}\cdot\sqrt{\frac{1}{I_n(p)}}$$

נציב

רווח בסמך:

$$\boxed{\frac{n}{\sum x_i}\pm Z_{\frac{\alpha}{2}}\cdot\sqrt{\frac{1}{\frac{\left(\sum^{n} x_i\right)^2}{n}-\frac{n-\sum x_i}{\left(1-\frac{n}{\sum x_i}\right)^2}}}}$$

2 א)

$$P\left(\frac{X-\mu}{\sigma} - \frac{(T-\mu)}{\sigma}\right) \Rightarrow T = \sigma Z_{0.95} + \mu$$

ב) נבנה

$$L(\sigma) = \prod_{i=1}^{n} f(x_i, \sigma) = \prod_{i=1}^{n} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \right)$$

$$l(\sigma) = \sum_{i=1}^{n} \log\left(\frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \right) =$$

$$-n\log(\sigma) - \frac{n}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x-\mu)^2$$

$$l'(\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x-\mu)^2 = 0 \iff \frac{1}{\sigma^2}\sum_{i=1}^{n}(x-\mu)^2 = n$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(x-\mu)^2}{n}}$$

כפי שראינו רווח סמך מקורב הינו $\hat{\sigma} \pm z_{\frac{\alpha}{2}} \hat{se}$, נחפש את $\hat{se}$

$$\hat{se} = \sqrt{\frac{1}{I_n(\sigma)}}$$

$$S(x,\sigma) = \frac{\partial}{\partial \sigma}\left(-\log(\sigma) - \log(\sqrt{2\pi}) - \frac{(x-\mu)^2}{2\sigma^2} \right) = \frac{-1}{\sigma} + \frac{(x-\mu)^2}{\sigma^3}$$

$$-S'(x,\sigma) = -\frac{1}{\sigma^2} + \frac{(x-\mu)^2 \cdot 3\sigma^2}{\sigma^6} = -\frac{1}{\sigma^2} + \frac{3(x-\mu)^2}{\sigma^4}$$

$$I(\sigma) = -\mathbb{E}[S'(x,\sigma)] = \mathbb{E}\left[-\frac{1}{\sigma^2} + \frac{3(x-\mu)^2}{\sigma^4} \right] = -\frac{1}{\sigma^2} + \frac{3}{\sigma^2}\mathbb{E}[Z^2] =$$

$$= -\frac{1}{\sigma^2} + \frac{3}{\sigma^2}\left[\mathrm{var}(z) + \overset{0}{\mathbb{E}}(z)^2 \right) = \frac{2}{\sigma^2} = I(\hat{\sigma})$$

רווח הסמך הינו

$$\hat{se} = \sqrt{\frac{1}{I_n(\hat{\sigma})}} = \sqrt{\frac{1}{n\frac{2}{\hat{\sigma}^2}}} \Rightarrow \boxed{\hat{\sigma} \pm Z_{0.025} \cdot \sqrt{\frac{\hat{\sigma}^2}{2n}}}$$

שאלה 2

ג)

$z_{0.95} \approx 2$

בסעיף א׳ אמרנו כי $J \approx 2\sigma + \mu$

ולכן האומ״נ יהיה $\hat{J} = 2\hat{\sigma} + \hat{\mu} = 2\sqrt{\Sigma_i (x_i - \bar{x}_n)^2} + \bar{X}_n$

מאחר שמתקיים $\frac{\hat{\sigma}_n - \sigma}{se(\sigma)} \approx N(0,1)$, $\hat{J} = g(\hat{\sigma}_n)$

מתוך שיטת הדלתא מקבל

$$N(0,1) \approx \frac{\hat{\tau} - \tau}{2\sqrt{\frac{\hat{\sigma}_n^2}{2n}}} = \frac{\hat{\tau} - \tau}{|g'(\sigma)| \cdot \hat{se}(\sigma)}$$

ומכאן רווח הסמך הוא $\hat{\tau} \pm 4\sqrt{\frac{\hat{\sigma}_n^2}{2n}}$

כלומר

$$2\hat{\sigma}_n + \hat{\mu} \pm 4\sqrt{\frac{\hat{\sigma}_n^2}{2n}}$$

ד) שוב ניתן להעזר בשיטת הדלתא, מתקיים $\psi' = \frac{1}{\sigma}$

אני ניתן לגזור שהאומ״נ הוא $\hat{\psi}_n = \log(\hat{\sigma}_n)$ ומכאן גם

נמצא את רווח הסמך $\hat{\psi} \pm 2\frac{1}{\hat{\sigma}}\sqrt{\frac{\hat{\sigma}_n^2}{2n}}$

שאלה 4

א)

$$B(\theta)=P_\theta(X_{(n)}>c)=1-P(X_n\le c)=1-\prod_i^n p(X_i\le c)$$

$$\begin{cases} 0 & c>\theta \\ 1 & c<0 \\ 1-\left(\frac{c}{\theta}\right)^n & else \end{cases}$$

ב)

$$\alpha:=\sup_{\theta\in\Theta_0} B(\theta)\overset{\substack{\text{מונוטוני, מקסימום}\\ \text{עבור } \theta=\frac12}}{\le} B\left(\tfrac12\right)=\begin{cases} 1-(2c)^n, & \theta\ge c \\ 0, & \theta<c\end{cases}$$

ג) נחפש c כך שגודל המבחן יהיה שווה ל 0.05

$$1-(2c)^n\le 0.05$$
$$\Updownarrow$$
$$0.95\le(2c)^n$$
$$\Updownarrow$$
$$\boxed{\frac{\sqrt[n]{0.95}}{2}\le c}\Longrightarrow$$

נבחר $\frac{\sqrt[n]{0.95}}{2}=c$

3) $0.498=\frac{\sqrt[20]{0.95}}{2}=c \Leftarrow n=20$

לפי שאלה 3.

$$p\text{-value}=P_{0.5}(X_{(20)}\ge 0.48)=1-\left(\frac{0.48}{0.5}\right)^{20}=0.558>c$$

ולכן לא נדחה את השערת האפס ברמת מובהקות 0.05.

שאלה 6

א) אנו יודעים כי האומד הנ"ל אסמפטוטי נורמלי $\hat{\mu}=\bar{X}_n$

$$-\log f(x,\mu) = -\log\left(\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}\right) = \log(\sigma)+\frac{1}{2}\log(2\pi)+\left(\frac{x-\mu}{2}\right)^2\cdot\frac{1}{\sigma^2}$$

$$-(\log f(x,\mu))' = \frac{x-\mu}{\sigma^2} \qquad -(\log f(x,\mu))'' = \frac{1}{\sigma^2}$$

$$\Downarrow$$

$$I(\hat{\mu}) = -E\left[\frac{\partial^2}{\partial^2\mu}\log f(x,\mu)\right] \Rightarrow W = \frac{\hat{\mu}-\mu_0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

נדחה אמ"מ $Z_{\frac{\alpha}{2}} < |W|$

ב) שוב מאחר ושהאומד הוא נורמלי אסמפטוטי, כפי שראינו בשאלה 2

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

לפי שיטת הדלתא, $\hat{se}(\hat{\sigma}^2) = 2\hat{\sigma}_n\cdot\hat{se}(\hat{\sigma}_n) =$

$$= 2\hat{\sigma}_n^2\cdot\frac{1}{\sqrt{2n}} = \hat{\sigma}_n^2\cdot\sqrt{\frac{2}{n}}$$

ומאחר הדבר נקבל כי $W = \frac{\hat{\sigma}^2-\sigma_0^2}{\hat{\sigma}^2\cdot\sqrt{\frac{2}{n}}}$

כאשר $W > Z_{\frac{\alpha}{2}}$ נדחה את המבחן.