תרגול כיתה – רווחי סמך ובדיקת השערות

רווח סמך

זוהי אמידה מרווחית (באינטרוול) של פרמטר באוכלוסייה, ע"י שימוש בתוצאות מדגם. רווח הסמך נבנה כך שההסתברות שהפרמטר יהיה בתוך הקטע נקבעת מראש ושווה $1-\alpha$. הסתברות זו נקראת "רמת סמך" או "רמת בטחון".

α נקראת "רמת מובהקות" (היא השגיאה המותרת באומדן) $0 \le \alpha \le 1$.

<u>רווח סמך עבור *תוחלת* האוכלוסייה, כאשר *שונותה ידועה*</u>

רו"ס ברמת סמך $1-\alpha$ עבור μ: $\bar{X} \pm z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$

<u>רווח סמך עבור *תוחלת* האוכלוסייה, כאשר *שונותה אינה ידועה*</u>

נסמן את סטיית התקן המדגמית: $S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum x_i^2 - n\bar{x}^2}{n-1}}$

רו"ס ברמת סמך $1-\alpha$ עבור μ: $\bar{X} \pm t_{\left(n-1,1-\frac{\alpha}{2}\right)} \frac{S}{\sqrt{n}}$

שאלה 1

מכונה מייצרת ברגים שאורכם מתפלג נורמלית עם תוחלת 4 ס"מ וסטיית תקן 0.2 ס"מ.
עקב תקלה במכונה הועלה חשד כי המכונה אינה מייצרת ברגים באורך הנדרש.
לשם בדיקה נלקח מדגם מקרי של 25 ברגים אשר יוצרו במכונה לאחר גילוי התקלה ונמצא כי האורך הממוצע שלהם הוא 3.9 ס"מ . נניח כי השונות נותרה ללא שינוי.

א. מהו רווח סמך (רו"ס) לתוחלת אורך הברגים לאחר התקלה ברמת מובהקות (ר"מ) 2% ?
ב. רוצים למצוא רווח סמך לתוחלת, ברמת בטחון של 95%, ע"י לקיחת מדגם של מספר מסויים של ברגים. כמה ברגים לפחות צריך לקחת על מנת שאורכו של רווח הסמך יהיה לכל היותר 0.1 ס"מ ?
ג. בנו רו"ס בגודל 0.118 ס"מ בעזרת מדגם בגודל 49 ברגים. מה רמת המובהקות α ?

<u>פתרון:</u>

א). נסמן ב- X את אורך הברגים לאחר התקלה , $X \sim N(\mu, 0.2^2)$. הפרמטר μ – לא ידוע.

כמו כן נתון : $n = 25$ ו- $\bar{X} = 3.9$

ונתונה רמת מובהקות של 2% $\Leftarrow$ $\alpha = 0.02$ $\Leftarrow$ $Z_{1-\frac{\alpha}{2}} = 2.326$

רו"ס ברמת מובהקות של α לתוחלת כאשר השונות ידועה:

$$\bar{X} - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

נציב את הנתונים ונקבל: $3.9 - 2.236 \cdot \frac{0.2}{5} \le \mu \le 3.9 + 2.236 \cdot \frac{0.2}{5} \Rightarrow \boxed{3.81 \le \mu \le 3.99}$

ב). נשתמש בנוסחה לרו"ס ברמת מובהקות של α לתוחלת כאשר השונות ידועה:

$$\bar{X}-Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}\le\mu\le\bar{X}+Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$$

מכאן שגודל רווח הסמך הוא: $L=2Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$

במקרה שלנו, רמת ביטחון של 95% $\Leftarrow$ $\alpha=0.05$ $\Leftarrow$ $Z_{1-\frac{\alpha}{2}}=1.96$.

נרצה שאורך רווח הסמך יהיה לכל היותר 0.1 לכן:

$$2\cdot1.96\cdot\frac{0.2}{\sqrt{n}}\le0.1\Rightarrow\sqrt{n}\ge7.84\Rightarrow n\approx61.47.$$

ז"א שיש צורך במדגם בגודל 62 ברגים לפחות.

ג). גודל רווח הסמך הוא: $L=2Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$

במקרה שלנו צריך למצוא $\alpha=?$

$$2Z_{1-\frac{\alpha}{2}}\cdot\frac{0.2}{\sqrt{49}}=0.118\quad\Rightarrow Z_{1-\frac{\alpha}{2}}=2.06\Rightarrow1-\frac{\alpha}{2}=0.98\Rightarrow\alpha=0.04$$

שאלה 2

חוקר מציע סוג של דשן שעשוי להגדיל את משקלם של תפוחים. נלקח מדגם מקרי של 9 תפוחים והתקבלו התוצאות הבאות (בגרמים): $\sum x=500,\ \sum x^2=28166$

א. אמוד ברווח סמך את משקל התפוחים, ברמת סמך של 90%, על סמך תוצאות המדגם הנ"ל.
ב. לאחר האמידה התברר שחלה תקלה שיטתית במדידות ויש להוסיף 8 גרם לכל ערך שנמדד. הסבר בקצרה, האם וכיצד ישפיע תיקון הטעות על גודל ומיקום רווח הסמך.

פתרון:

א). נחשב את ממוצע המדגם וסטיית התקן המדגמית:

$$\bar{x}=\frac{\sum x_i}{n}=\frac{500}{9}=55.556,\qquad S=\sqrt{\frac{\sum x_i^2-n\bar{x}^2}{n-1}}=\sqrt{\frac{28166-9\cdot55.556^2}{9-1}}=6.966$$

הערך המתאים לנתוני השאלה – מטבלת t:

$n=9,\quad 1-\alpha=0.9$ <= נמצא בטבלה: $t_{8,0.95}=1.86$

נציב בנוסחת רו"ס:

$$55.55-1.86\cdot\frac{6.966}{3}\le\mu\le55.55+1.86\cdot\frac{6.966}{3}$$

$$\boxed{51.231\le\mu\le59.868}$$

ב). השפעת תיקון הטעות:

הוספה של קבוע לא משפיעה על השונות (וסטיית התקן), לכן גודל רווח הסמך לא ישתנה.

הוספה של קבוע מעלה את הממוצע בקבוע, כלומר הממוצע יגדל ב-8. לכן מיקום רו"ס יוזז ב-8 יח' ימינה.

<u>רווח סמך לשונות האוכלוסייה</u>

האומד לשונות האוכלוסייה הוא: S^2

רווח סמך ברמת סמך $1-\alpha$ עבור שונות האוכלוסייה σ^2 :

$$\boxed{\frac{(n-1)\cdot S^2}{\chi^2_{(n-1);1-\alpha/2}} \le \sigma^2 \le \frac{(n-1)\cdot S^2}{\chi^2_{(n-1);\alpha/2}}}$$

<u>הערה 1</u>: מבחינה טכנית נוח להשתמש בנוסחה $(n-1)\cdot S^2 = \sum_{i=1}^{n}(x_i-\bar{X})^2$.

<u>הערה 2</u>: אם נתונה <u>תוחלת</u> האוכלוסייה (μ), המבחן נשאר כנ"ל, עם 2 השינויים הבאים:

(1) יש להשתמש ב- μ במקום ב- $\bar{X}$. (2) דרגות החופש משתנות ל- n במקום $(n-1)$.

<u>שאלה 3</u>

נלקח מדגם מקרי של 5 אנשים שניגשו למבחן IQ וקיבלו את הציונים הבאים: 105 ,104 ,102 ,98 ,96.

מצא ברמת בטחון 99% -

א. רווח סמך לשונות ורווח סמך לסטיית תקן של הציונים.

ב. מצא רווח סמך לשונות הציונים כאשר מניחים שהתוחלת ידועה ושווה ל- 101.

<u>פתרון</u>:

מהנתונים: $n=5, \quad \bar{X}=\frac{1}{5}\sum_{i=1}^{5}x_i=101, \quad \sum_{i=1}^{5}(x_i-\bar{X})^2=60$

רמת הבטחון 99% לכן: $1-\alpha=0.99 \Rightarrow \alpha=0.01 \quad \Rightarrow \alpha/2=0.005$

מטבלת חי-בריבוע:

$$\chi^2_{(n-1);\alpha/2}=\chi^2_{4;0.005}=0.207 \qquad \chi^2_{(n-1);1-\alpha/2}=\chi^2_{4;0.995}=14.86$$

א). רווח הסמך לשונות:

$$\frac{60}{14.86} \le \sigma^2 \le \frac{60}{0.207} \quad \Rightarrow \boxed{4.0376 \le \sigma^2 \le 289.855}$$

רווח הסמך לסטיית התקן:

$$\sqrt{4.0376} \le \sigma \le \sqrt{289.855} \quad \Rightarrow \boxed{2.01 \le \sigma \le 17.025}$$

ב). כעת נציב את התוחלת הידועה $\mu=101$: $\sum_{i=1}^{5}(x_i-\mu)^2=60$

מטבלת חי-בריבוע:

$$\chi^2_{n;\alpha/2}=\chi^2_{5;0.005}=0.412 \qquad \chi^2_{n;1-\alpha/2}=\chi^2_{5;0.995}=16.75$$

רווח הסמך לשונות:

$$\frac{60}{16.75} \le \sigma^2 \le \frac{60}{0.412} \quad \Rightarrow \boxed{3.582 \le \sigma^2 \le 145.631}$$

הערה: בגלל שנתונה לנו התוחלת במדוייק ולא נאלצנו לאמוד אותה, רווח הסמך קטן יותר מאשר בסעיף א'.

<u>בדיקת השערות</u>

בבדיקת השערות מתחילים בהשערה סטטיסטית על פרמטר באוכלוסייה.

קיימות 2 השערות משלימות האחת לשניה לגבי פרמטר לא ידוע באוכלוסייה, ורוצים לקבוע מי מהשתיים נכונה.

נסמן:

H_0 = השערת האפס, H_1 = ההשערה האלטרנטיבית.

השערת האפס = מייצגת את המצב הקיים.

ההשערה האלטרנטיבית = מייצגת את המצב החדש.

דחיית השערה אחת משמעותה קבלת ההשערה האחרת.

כדי לדעת האם לדחות או לקבל את השערת האפס נחלק את תחום ערכי ההתפלגות ל- 2 חלקים (לאו דווקא שווים), חלק אחד נקרא אזור הדחייה של H_0 והחלק השני נקרא אזור הקבלה של H_0.

הנקודה המפרידה בין השניים היא הנקודה הקריטית (קרוי גם: הערך הקריטי) K.

הפרמטרים הנקבעים ע"י החוקר: רמת מובהקות - α; עוצמת המבחן - $1-\beta$.

<u>השלבים בבדיקת השערות</u>:

1. ניסוח ההשערות.
2. מדידת הסטטיסטי המופיע בהשערה על ידי המדגם (אומד).
3. הקצאת רמת מובהקות (לדוגמה: 1%, 5%, 10%).
4. קביעת אזורי קבלה ואזורי דחיה של H_0 על סמך רמת המובהקות (הנתונה).
5. בדיקה אם האומד נמצא באזור הקבלה או הדחיה של H_0 והחלטה בהתאם.

<u>טעויות בבדיקת השערות</u>

החלטה \ מציאות	H_0 נכונה	H_1 נכונה
קבלת H_0	<u>החלטה נכונה</u>	טעות מסוג שני β
קבלת H_1	טעות מסוג ראשון α (רמת מובהקות)	<u>החלטה נכונה</u> (עוצמת המבחן = $\pi = 1-\beta$)

<u>הקשר בין רווח סמך ובדיקת השערות</u>

רווח סמך ברמת בטחון $(1-\alpha)$ הוא גם איזור הקבלה של מבחן השערות דו-כיווני ברמת מובהקות α.

<u>בדיקת השערות עבור *תוחלת* האוכלוסייה, כאשר *שונותה ידועה*</u>

<u>בדיקת השערות</u>: נסמן- μ_0 – ממוצע האוכלוסייה הקיים (השערת האפס).

מבחן ההשערות: $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu >,<,\neq \mu_0 \end{cases}$

סטטיסטי המבחן: $Z \sim \dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ לחילופין, הערך הקריטי של המבחן: $K = \mu_0 + Z_{1-\alpha} \cdot \dfrac{\sigma}{\sqrt{n}}$.

בדיקת השערות עבור *תוחלת* האוכלוסייה, כאשר *שונותה איננה ידועה*

נסמן את סטיית התקן המדגמית: $S=\sqrt{\dfrac{\sum\left(x_i-\bar{x}\right)^2}{n-1}}=\sqrt{\dfrac{\sum x_i^2-n\bar{x}^2}{n-1}}$

<u>בדיקת השערות</u>: נסמן- μ_0 – ממוצע האוכלוסייה הקיים (השערת האפס).

סטטיסטי המבחן: $\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}\sim t_{n-1}$ לחילופין, הערך הקריטי של המבחן: $K=\mu_0+t_{n-1}\cdot\dfrac{S}{\sqrt{n}}$.

<u>*P-Value* (P.V.)</u>

זוהי רמת המובהקות המינימלית (ה- α המינימלית) עבורה נדחה את H_0.

כלומר אם: $\alpha <$ *p-value* <= אזי נדחה את H_0, אחרת לא נדחה את H_0 (נקבל את H_1).

<u>שאלה 1</u>

בנבחרת ריצה הזמן הממוצע של הרצים במירוץ 60 מ', הוא 7.6 שניות עם סטיית תקן של 1.4.
המאמן מציע שיטת אימון חדשה. השיטה החדשה נבדקה על מדגם בגודל 16 אצנים והתקבל שזמן הריצה הממוצע ירד ל- 6.9 שניות.

א. בדוק ברמת מובהקות של 5% האם שיטת האימון החדשה עדיפה על הקיימת.
ב. חשב את ערך *p-value* של נתוני המדגם.

<u>פתרון</u>:

א). צריך לבצע מבחן חד-כיווני שמאלי לכך שממוצע האוכלוסייה (נבחרת הריצה) ירד.

הנתונים: $\mu_0=7.6,\ \sigma=1.4,\ n=16,\ \bar{X}=6.9,\ \alpha=0.05$

השערות המבחן: $\begin{cases}H_0: \mu=7.6\\ H_1: \mu<7.6\end{cases}$

נמצא את הערך הקריטי: $K=\mu_0-Z_{1-\alpha}\cdot\dfrac{\sigma}{\sqrt{n}}=7.6-1.645\cdot\dfrac{1.4}{4}=7.02425$

מכיוון שמתקיים- $\bar{X}=6.9<7.02425$ <= אזי דוחים את H_0 בר"מ 5%.

קיבלנו שברמת מובהקות של 5% השערת האפס תידחה ,כלומר ניתן להסיק ששיטת האימון החדשה עדיפה על שיטת האימון הנוכחית (מקבלים H_1).

ב). חישוב ה- *p-value*:

$$\text{P.V.}=P(\bar{X}_{16}\le 6.9)=P\left(Z\le\frac{6.9-7.6}{1.4/\sqrt{16}}\right)=P(Z\le -2.0)=1-P(Z\le 2.0)=0.0228$$

ניתן לראות לפי ערך ה- *p-value* כי $(\alpha=0.05)<(p\text{-}value=0.0228)$,

לכן, כפי שקיבלנו בסעיף א', ההשערה תידחה עבור $\alpha=5\%=0.05$ הנתון.

שאלה 2

להלן נתונים על משקלם (בגרמים) של 12 עכברים אשר הואכלו במשך 28 ימים בשמן דגים:

24, 23.5, 24, 24, 25, 22.5, 20, 23.5, 28.5, 18, 20, 26

ידוע שמשקל ממוצע של עכבר רגיל מסוג זה הוא 22 גרם. התייחס למשקל ממוצע זה כאל התוחלת.

רוצים לבדוק את ההשערה ששמן דגים משפיע באופן מובהק על משקל העכבר הממוצע ברמת מובהקות $\alpha = 0.05$. נסח ובצע את המבחן והסק את המסקנות.

<u>פתרון</u>:

ברצוננו לבדוק את ההשערות:

$$\begin{cases} H_0 : \mu = 22 \\ H_1 : \mu \neq 22 \end{cases}$$

כאשר השונות לא ידועה.

נשתמש במבחן דו צדדי, אנו מבקשים לבדוק קיום השפעה שכיוונה לא ידוע לנו:

דחה H_0 אם $\bar{x} < \mu_0 - t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$ -או- $\bar{x} > \mu_0 + t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$.

נחשב: $S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = 2.848$ $\bar{x} = \frac{\sum x_i}{n} = 23.25$.

הערך המתאים מטבלת t: $(1-\alpha = 0.95,\ n = 12)$ <= $t_{12-1,1-\alpha/2} = t_{11,0.975} = 2.201$

נחשב את הערך הקריטי K_1 של המבחן:

$$K_1 = \mu_0 + t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} = 22 + 2.201 \cdot \frac{2.848}{\sqrt{12}} = \boxed{23.81}$$

בכיוון השני הערך הקריטי K_2 הוא: $K_2 = 22 - 2.201 \cdot 2.848 / \sqrt{12} = \boxed{20.19}$

רואים ש- $K_2 = 20.19 < \bar{x} = \boxed{23.25} < K_1 = 23.81$

כלומר הממוצע אינו חורג מתחום הערכים הקריטיים <= לכן <u>לא נדחה</u> את השערת האפס.

המסקנה: האכלה בשמן דגים לא משנה באופן מובהק (סטטיסטית) את משקל העכברים.