תרגול כיתה 3 – מבוא להסתברות וסטטיסטיקה

הסקה סטטיסטית – בדיקת השערות

בדיקת השערות

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 השנייה.
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>

מציאות / החלטה	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 = $\pi = 1-\beta$)

הקשר בין רווח סמך ובדיקת השערות

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

בדיקת השערות עבור *תוחלת* האוכלוסיה, כאשר *שונותה ידועה*

<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: $Z \sim \dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$

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

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: $\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}\sim t_{n-1}$

<u>בדיקת השערות לפרופורציה באוכלוסיה</u>

p – פרופורציית התכונה באוכלוסייה. $\hat{p}=x/n$ – פרופורציה מדגמית.

<u>בדיקת השערות</u>: עבור $H_0: p=p_0$ – סטטיסטי המבחן: $Z=\dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$

<u>*P-Value*</u> (*P.V.*) היא רמת המובהקות המינימלית (ה- α המינימלית) עבורה נדחה את H_0.

כלומר, אם: α > *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* מנתוני 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: $\dfrac{K-\mu_0}{\sigma/\sqrt{n}}\sim Z_{1-\alpha}\Rightarrow K=\mu_0-Z_{1-\alpha}\cdot\dfrac{\sigma}{\sqrt{n}}=7.6-1.645\cdot\dfrac{1.4}{4}=7.02425$

מכיוון שמתקיים- $\bar{X}=6.9<7.02425$ <= לכן דוחים 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*:

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

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.

<u>שאלה 2</u>

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.

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 שלא ידוע לנו כיוונה:

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

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

נבדוק האם- $\boxed{23.81} = 22 + 2.201 \cdot \frac{2.848}{\sqrt{12}} = \mu_0 + t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} \overset{(?)}{<} \bar{x} = \boxed{23.25}$.

הכיוון השני: $\boxed{23.25} = \bar{x} \overset{(?)}{<} 22 - 2.201 \cdot 2.848 / \sqrt{12} = \boxed{20.19}$

אי השוויון לא מתקיים <= ולכן <u>לא נדחה</u> את השערת האפס.

(הערה: במקרה זה די ברור שגם הכיוון השני לא מתקיים, אבל במקרה של ספק – חובה לבדוק).

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

<u>שאלה 3</u>

מניסיון העבר ידוע כי ההסתברות לנביטת זרע של חיטה אפריקאית היא 80%. חוקר מציע שיטה חדש שלטענתו עשויה להעלות את אחוז הנביטה ל- 90%. הוחלט להפעיל את השיטה על מדגם מקרי של 50 זרעים. בתום הניסוי התברר כי 42 זרעים נבטו.

בדוק בר"מ 5% האם שיטת החוקר אכן יעילה.

פתרון:

$X \sim Bin(50, p)$ – מספר הזרעים שנבטו.

$$p_0 = 0.8, \qquad p_1 = 0.9, \qquad n = 50, \qquad \hat{p} = 42/50 = 0.84$$

השערות המבחן החד-צדדי ימני לפרופורציה:

$$\begin{cases} H_0 : p = 0.8 \\ H_1 : p > 0.8 \end{cases}$$

$$K = p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} = 0.8 + 1.65 \cdot \sqrt{\frac{0.8 \cdot 0.2}{50}} = 0.893$$

$\hat{p} = 0.84 < 0.893 \quad \Leftarrow$ לכן לא דוחים את H_0 (מקבלים את השערת האפס) בר"מ 5% <=
המסקנה ששיטת החוקר איננה יעילה.

שאלה 4

מניסיון העבר ידוע שזמן התגובה של נהג למכשול בכביש מפולג נורמלית עם תוחלת של 1.4 שניות וסטיית תקן של 0.24 שניות.

הועלתה טענה שתרופה חדשה פוגעת בערנות הנהג ומעלה את זמן תגובתו.

הוחלט שלמדגם מקרי של 25 נהגים תינתן התרופה ויימדד ממוצע זמן תגובתם.

חוקר שנשכר לבדיקת הטענה נוקט בכלל ההחלטה הבא:

"התרופה תפסל לשימוש אם ממוצע זמן התגובה במדגם יעלה על 1.5 שניות".

א. מהי ההסתברות לטעות מסוג I לפי הקריטריון של החוקר?
ב. מהי ההסתברות לטעות מסוג II אם התרופה אכן מעלה את זמן התגובה ל- 1.54 שניות?
ג. מהי עוצמת המבחן אם התרופה אכן מעלה את זמן התגובה ל- 1.54 שניות?

פתרון:

א). הערך הקריטי שקבע החוקר הוא $C = 1.5$.

נחשב את סטטיסטי המבחן $Z_{\bar{X}} = \frac{1.5 - 1.4}{0.24/\sqrt{25}} = 2.0833$

מהטבלה: $\Phi(2.0833) = 0.981$.

לכן ההסתברות לטעות מסוג I: $\alpha = 1 - 0.981 = 0.019$

ב). שוב נחשב את הסטטיסטי, הפעם תחת ההנחה $\mu = 1.54$

$$Z_{\bar{X}} = \frac{1.5 - 1.54}{0.24/\sqrt{25}} = -0.833$$

מהטבלה: $\Phi(-0.833) = 0.202$.

לכן ההסתברות לטעות מסוג II: $\beta = 0.202$

ג). עוצמת המבחן $power = 1 - \beta$.

לכן: $power = 1 = \beta = 1 - 0.202 = 0.798$