סטטיסטיקה / תרגיל 9#

אריאל סטולרמן
קבוצה 03

(1)

(a) – (n) : נפתרו בתרגול.

(o)

אם האלטרנטיבה היתה $H_1: \mu \neq 2$, מובהקות התוצאה היתה מחצית מאשר קודם :

$$\frac{1}{2} \cdot p - value = P_{H_0}(\overline{X_4} > 2.45) = 0.0359 \Rightarrow \boldsymbol{p - value = 0.0718}$$

(p)

נתון : $H_1: \overline{X_3} \sim N(2.8, \frac{0.5}{\sqrt{3}})$

- $\text{עוצמת המבחן} = \pi = P_{H_1}(\overline{X_3} > 2.5) = 1 - \Phi\left(\frac{2.5-2.8}{\frac{0.5}{\sqrt{3}}}\right) = 1 - \Phi(-1.039) = 0.85$
- $P(\text{טעות מסוג ראשון}) = P_{H_0}(\overline{X_3} > 2.5) = 1 - \Phi\left(\frac{2.5-2}{\frac{0.5}{\sqrt{3}}}\right) = 0.041$ (לא השתנה)
- $P(\text{טעות מסוג שני}) = 1 - \pi = 0.15$

(q)

עוצמת המבחן **קטנה** ואילו טעות מסוג ראשון **נשארה אותו דבר**.
יכולנו לצפות שעוצמת המבחן תקטן : קירבנו את תוחלת האלטרנטיבה "שמאלה" על הציר לכיוון הערך הקריטי, מה שהביא ל"ספירת" פחות מסה תחת האלטרנטיבה החדשה לעומת הישנה.
יכולנו לצפות את השפעת השינוי של H_1 על טעות מסוג ראשון גם ללא חישובים כיוון שברור היה שערך זה לא ישתנה, מפני שטעות מסוג ראשון אינה תלויה כלל באלטרנטיבה (רק ב-H_0 ובמבחן).

(r) להלן הקוד לבדיקת ההשערות באמצעות סימולציה, וניתן לראות כי קיבלנו תוצאה **זהה** לסעיף (g) :

```
> sigma=0.5
> mu=2
> x1=rnorm(1000,mu,sigma)
> x2=rnorm(1000,mu,sigma)
> x3=rnorm(1000,mu,sigma)
> xmean=(x1+x2+x3)/3
>
> p=0
> for(i in 1:1000){
+ if(xmean[i] > 2.5)(p <- p+1)
+ }
>
> p=p/1000
> p
[1] 0.041
> sigma=0.5
```

```
> mu=2.8
> x1=rnorm(1000,mu,sigma)
> x2=rnorm(1000,mu,sigma)
> x3=rnorm(1000,mu,sigma)
> xmean=(x1+x2+x3)/3
>
> p=0
> for(i in 1:1000){
+ if(xmean[i] > 2.5)(p <- p+1)
+ }
>
> p=p/1000
> p
[1] 0.844
```

בחישוב האחרון חישבנו את עוצמת המבחן, ולפיכך ב-15.6% מהמדגמים לא נצליח לזהות שזהו שרת מסוג ב', כלומר:
$\beta = 1 - \pi = 1 - 0.844 = 0.156$.

(2)

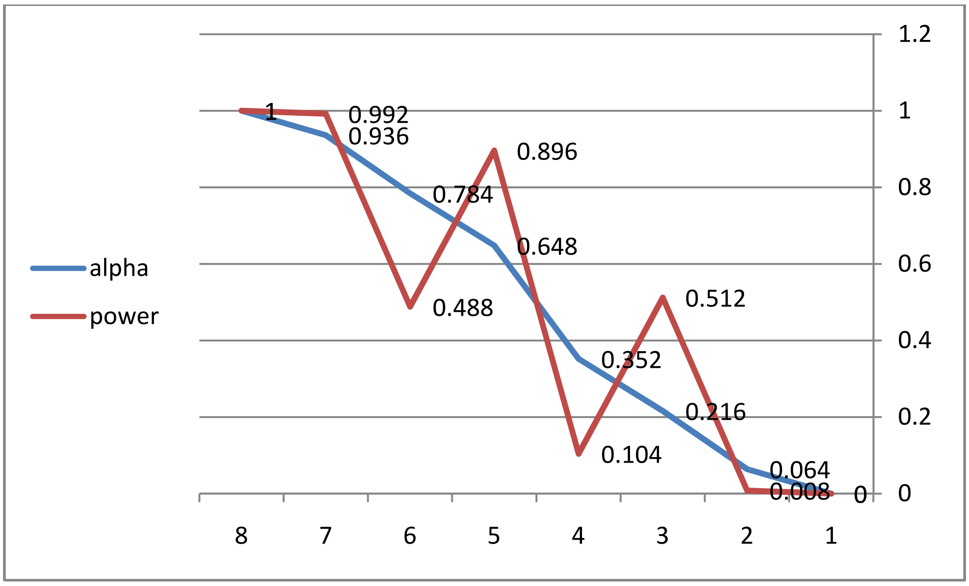
נסמן ב-p את הסיכוי לעץ.

(*a*) $H_0: p = 0.4$, $H_1: p = 0.2$

(*b,c*) מספר העצים הכולל שאנו יכולים לקבל הוא: *0,1,2,3*. להלן איזורי הדחיה השונים (הרציפים), ומשמעותם:

אזור דחיה	מספר העצים עבורו נדחה	הסיכוי לטעות מסוג ראשון	הסיכוי לטעות מסוג שני	עוצמת המבחן
$C_0 = (3, \infty)$	$\emptyset$	0	1	0
$C_1 = [0, \infty)$	$\{0,1,2,3\}$	1	0	1
$C_2 = [1, \infty)$	$\{1,2,3\}$	$\sum_{i=1}^{3}\binom{3}{i}\cdot\left(\frac{2}{5}\right)^{i}\cdot\left(\frac{3}{5}\right)^{3-i} = 0.784$	$\left(1-\frac{1}{5}\right)^{3} = 0.512$	$1 - 0.512 = 0.488$
$C_3 = [2, \infty)$	$\{2,3\}$	$\sum_{i=2}^{3}\binom{3}{i}\cdot\left(\frac{2}{5}\right)^{i}\cdot\left(\frac{3}{5}\right)^{3-i} = 0.352$	$\sum_{i=0}^{1}\binom{3}{i}\cdot\left(\frac{1}{5}\right)^{i}\cdot\left(\frac{4}{5}\right)^{3-i} = 0.896$	$1 - 0.896 = 0.104$
$C_4 = [3, \infty)$	$\{3\}$	$\left(\frac{2}{5}\right)^{3} = 0.064$	$\sum_{i=0}^{2}\binom{3}{i}\cdot\left(\frac{1}{5}\right)^{i}\cdot\left(\frac{4}{5}\right)^{3-i} = 0.992$	$1 - 0.992 = 0.008$
$C_5 = (-\infty, 0]$	$\{0\}$	$\left(1-\frac{2}{5}\right)^{3} = 0.216$	$\sum_{i=1}^{3}\binom{3}{i}\cdot\left(\frac{1}{5}\right)^{i}\cdot\left(\frac{4}{5}\right)^{3-i} = 0.488$	$1 - 0.488 = 0.512$
$C_6 = (-\infty, 1]$	$\{0,1\}$	$\sum_{i=0}^{1}\binom{3}{i}\cdot\left(\frac{2}{5}\right)^{i}\cdot\left(\frac{3}{5}\right)^{3-i} = 0.648$	$\sum_{i=2}^{3}\binom{3}{i}\cdot\left(\frac{1}{5}\right)^{i}\cdot\left(\frac{4}{5}\right)^{3-i} = 0.104$	$1 - 0.104 = 0.896$
$C_7 = (-\infty, 2]$	$\{0,1,2\}$	$1-\left(\frac{2}{5}\right)^{3} = 0.936$	$\left(\frac{1}{5}\right)^{3} = 0.008$	$1 - 0.008 = 0.992$

(*d*) להלן תרשים של העוצמה כנגד טעות מסוג ראשון:



(*e*) ניתן לפסול על הסף את המבחנים בהם עוצמת המבחן נמוכה מהסיכוי לטעות מסוג ראשון, שהם אזורי הדחיה: C_4, C_3, C_2.

(3)

(*a*) אם נתון כי עבור מבחן סטטיסטי כלשהו 0.04 =*p-value*, ברמת ביטחון של $\alpha = 0.05$ אנו **נידחה** את השערת האפס.

(*b*) במצב הנתון בשאלה **לא יתכן** שנגיע למסקנות שונות בשני המדגמים עבור אותה רמת ביטחון α.

(*c*) במקרה זה **יתכן** ונגיע למסקנות שונות עבור שני המדגמים. דוגמא: נניח ואנו מסתכלים על המקרה הנורמלי, ונגדיר את סטטיסטי המבחן שלנו להיות התוצאה הראשונה במדגם, ונניח ובשני המדגמים בשאלה קיבלנו כי התוצאה הראשונה (עבור הראשון – היחידה) היא אותה תוצאה, נסמנה q. במקרה זה, חישוב ערך ה-p עבור שני המדגמים יהיה שונה, כיוון שישנה תלות בגודל המדגם:

- במדגם הראשון (תצפית בודדת): $\alpha_1{}^*(q) = 1 - \Phi\left(\frac{q-\mu_0}{\sigma}\right)$
- במדגם השני (100,000 תצפיות): $\alpha_2{}^*(q) = 1 - \Phi\left(\frac{q-\mu_0}{\sigma/\sqrt{100{,}000}}\right)$

במקרה זה, אם נקבל רמת ביטחון $\alpha_2^* < \alpha < \alpha_1^*$, נדחה את H_0 עבור המדגם הראשון ולא נדחה עבור השני.

(*d*) לחישוב סטטיסטי המבחן **אין** צורך בהנחות על המודל ההסתברותי.

(*e*) לחישוב מובהקות התוצאה, כלומר ערך ה-p, **ישנו צורך** בהנחות על המודל ההסתברותי ממנו מגיעים הנתונים, כיוון שהחישוב כולל שימוש בפונקציית הסתברות על הנתונים הנקבעת כמובן ע"י סוג הנתונים ולא רק ע"י ערכם.

(*f*) סטטיסטי המבחן אינו יחיד גם לאחר ניסוח ההשערות במונחים סטטיסטיים (ישנו סטטיסטי מבחן בעל עוצמה מקסימלית, אך הוא לא הסטטיסטי היחיד כמובן).

(*g*) את הסיכוי לטעות מסוג ראשון ניתן לחשב ללא ידיעת האלטרנטיבה, היות והחישוב לא תלוי בה. לעומת זאת, לא ניתן לחשב ללא ידיעת השערת האפס, היות והחישוב הוא חישוב הסתברות תחת השערת האפס.

(*h*) באופן הפוך, במקרה של חישוב עוצמה ישנו צורך בידיעת האלטרנטיבה (חישוב הסתברות תחת האלטרנטיבה), ואין צורך בידיעת השערת האפס.

(i) נניח כי הנחת היסוד היא שאדם זכאי עד שהוכח אחרת. תחת הנחה זו:

i. הסיכוי שאדם אשם נמצא זכאי מתאים לטעות מסוג שני (β).

ii. הסיכוי שאדם זכאי נמצא אשם מתאים לטעות מסוג ראשון (α).

iii. הסיכוי שאדם אשם אכן נמצא אשם מתאים לעוצמת מבחן ($\pi = 1 - \beta$).

iv. רמת הביטחון שדורש שופט לפני ששולח אדם לכלא מתאים לרמת המובהקות (*p-value*).

(j) במקרה זה יתקיים:

$$P_{H_0}(S(X) \in C_\alpha) = \alpha, \qquad \beta = P_{H_1}(S(X) \notin C_\alpha) = P_{H_0}(S(X) \notin C_\alpha) = 1 - \alpha$$
$$\Rightarrow \pi = 1 - \beta = 1 - (1 - \alpha) = \alpha$$

(4)

(a) סטטיסטי המבחן יהיה k = מספר האתרים אליהם ניסינו להיכנס (בין אם הצלחנו ובין אם לא). כיוון הדחיה יהיה ערכים נמוכים, שכן אז היחס בין מספר האתרים שלא תומכים בפיירפוקס לאלו שכן יגדל.

(b) נסמן את השערתו של אבנר 0 כ-H_0 ואת של אבנר 1 כ-H_1. נחשב את מספר האתרים המקסימלי בו נבקר ועדיין נדחה את H_0, עבור סיכוי לטעות של $\alpha = 0.2$:

$$P_{H_0}(k \le c) = 0.2 \Leftrightarrow \sum_{k=1}^{c}\left(1 - \frac{1}{10}\right)^{k-1} \cdot \frac{1}{10} = 0.2 \Leftrightarrow \sum_{k=1}^{c}\left(\frac{9}{10}\right)^{k-1} = \sum_{k=0}^{c-1}\left(\frac{9}{10}\right)^{k} = 2 \Leftrightarrow$$

$$\frac{1 - \left(\frac{9}{10}\right)^{c-1}}{1 - \frac{9}{10}} = 2 \Leftrightarrow 0.8 = \left(\frac{9}{10}\right)^{c-1} \Leftrightarrow 0.72 = \left(\frac{9}{10}\right)^{c} \Leftrightarrow c = \log_{\frac{9}{10}} 0.72 = 3.11$$

מכאן שהערך המקסימלי עבורו נדחה את H_0 ונחליט כי אבנר 1 צודק הוא $\boldsymbol{k = 3}$.

i. אם האתר השני לא תומך בפיירפוקס, נחליט כי **אבנר 1 צודק**.

ii. עוצמת המבחן:

$$\pi = P_{H_1}(k \le 3) = \sum_{k=1}^{3}\left(1 - \frac{1}{5}\right)^{k-1} \cdot \frac{1}{5} = \frac{1}{5}\sum_{k=1}^{3}\left(\frac{4}{5}\right)^{k-1} = \mathbf{0.488}$$

(c) נחשב את מספר האתרים המקסימלי בו נבקר ועדיין נדחה את H_0, עבור סיכוי לטעות של $\alpha = 0.05$:

$$P_{H_0}(k \le c) = 0.05 \Leftrightarrow \sum_{k=1}^{c}\left(1 - \frac{1}{10}\right)^{k-1} \cdot \frac{1}{10} = 0.05 \Leftrightarrow$$

$$\frac{1 - \left(\frac{9}{10}\right)^{c-1}}{1 - \frac{9}{10}} = 0.5 \Leftrightarrow 0.95 = \left(\frac{9}{10}\right)^{c-1} \Leftrightarrow 0.855 = \left(\frac{9}{10}\right)^{c} \Leftrightarrow c = \log_{\frac{9}{10}} 0.855 = 1.48$$

מכאן שהערך המקסימלי עבורו נדחה את H_0 ונחליט כי אבנר 1 צודק הוא $\boldsymbol{k = 1}$.

i. אם האתר השני לא תומך בפיירפוקס, נחליט כי **אבנר 0 צודק**.

ii. עוצמת המבחן:

$$\pi = P_{H_1}(k \le 1) = \frac{1}{5} = \mathbf{0.2}$$

(*d*) להלן מבחן יחס נראות למקרה כללי של $H_1: X\sim G(p_1), H_0: X\sim G(p_0)$:

$$\Lambda(k)=\frac{P_{H_1}(X=k)}{P_{H_0}(X=k)}=\frac{(1-p_1)^k\cdot p_1}{(1-p_0)^k\cdot p_0}=\frac{\boldsymbol{p_1}}{\boldsymbol{p_0}}\cdot\left(\frac{\boldsymbol{1-p_1}}{\boldsymbol{1-p_0}}\right)^{\boldsymbol{k}}$$

i. עבור $p_1<p_0$, ככל ש-k יגדל, $\Lambda(k)$ גם יגדל, ונדחה לערכים גבוהים של $\Lambda(k)$. עבור $p_1>p_0$, ככל ש-k יגדל, $\Lambda(k)$ יקטן. לפיכך גם כאן נדחה לערכים גבוהים של $\Lambda(k)$. כך או כך דוחים לערכים גבוהים של $\Lambda(k)$.

ii. אם $p_1<p_0$ נדחה עבור ערכים גבוהים של k. חישוב סף ההחלטה עבור הסיכוי לטעות α :

$$P_{H_0}(k>c)=\alpha$$

(*e*) $H_1: X\sim G(\frac{4}{5}), H_0: X\sim G\left(\frac{1}{5}\right), X$ מספר האתרים אליהם ניסו להיכנס.

(*f*) נחשב את מספר האתרים המקסימלי בו נבקר ועדיין נדחה את H_0, עבור סיכוי לטעות של $\alpha=0.05$:

$$P_{H_0}(k\le c)=0.05\Leftrightarrow\sum_{k=1}^{c}\left(1-\frac{1}{5}\right)^{k-1}\cdot\frac{1}{5}=0.05\Leftrightarrow\sum_{k=1}^{c}\left(\frac{4}{5}\right)^{k-1}=\sum_{k=0}^{c-1}\left(\frac{4}{5}\right)^{k}=0.25\Leftrightarrow$$

$$\frac{1-\left(\frac{4}{5}\right)^{c-1}}{1-\frac{4}{5}}=0.25\Leftrightarrow 0.95=\left(\frac{4}{5}\right)^{c-1}\Leftrightarrow 0.76=\left(\frac{4}{5}\right)^{c}\Leftrightarrow c=\log_{\frac{4}{5}}0.76=1.23$$

מכאן שהערך המקסימלי עבורו נדחה את H_0 ונחליט כי אבנר 1 צודק הוא $\boldsymbol{k=1}$.

i. אם האתר השני לא תומך בפיירפוקס, נחליט כי **אבנר 0 צודק**.

ii. עוצמת המבחן :

$$\pi=P_{H_1}(k\le 1)=\frac{4}{5}=\mathbf{0.8}$$

(*g*) נחשב את מספר האתרים המקסימלי בו נבקר ועדיין נדחה את H_0, עבור סיכוי לטעות של $\alpha=0.2$:

$$P_{H_0}(k\le c)=0.2\Leftrightarrow\sum_{k=1}^{c}\left(1-\frac{1}{5}\right)^{k-1}\cdot\frac{1}{5}=0.2\Leftrightarrow\sum_{k=1}^{c}\left(\frac{4}{5}\right)^{k-1}=\sum_{k=0}^{c-1}\left(\frac{4}{5}\right)^{k}=1\Leftrightarrow$$

$$\frac{1-\left(\frac{4}{5}\right)^{c-1}}{1-\frac{4}{5}}=1\Leftrightarrow 0.8=\left(\frac{4}{5}\right)^{c-1}\Leftrightarrow 0.64=\left(\frac{4}{5}\right)^{c}\Leftrightarrow c=\log_{\frac{4}{5}}0.64=2$$

מכאן שהערך המקסימלי עבורו נדחה את H_0 ונחליט כי אבנר 1 צודק הוא $\boldsymbol{k=2}$.

i. אם האתר השני לא תומך בפיירפוקס, נחליט כי **אבנר 1 צודק**.

ii. עוצמת המבחן :

$$\pi=P_{H_1}(k\le 2)=\sum_{k=1}^{2}\left(1-\frac{4}{5}\right)^{k-1}\cdot\frac{4}{5}=\frac{4}{5}\sum_{k=1}^{2}\left(\frac{1}{5}\right)^{k-1}=\mathbf{0.96}$$

iii. השינוי לא מפתיע בשל גדילת הפרמטרים – גם p_0 המשפיע על איזור הדחיה וגם p_1 המשפיע על העוצמה.

(5) (*a*), (*b*), (*f*) : נפתרו בתרגול.

(*c*) מבחן זה יהיה בעל עוצמה מקסימלית, כיוון שיחס הנראות מונוטוני במספר השאילתות.

(*d*) להתפלגות לא פואסונית ויחס נראות עדיין מונוטוני במספר השאילתות, המבחן עדיין יהיה בעל עוצמה מקסימלית.

(*e*) אם יחס הנראות לא יהיה מונוטוני במספר השאילתות, המבחן לא יהיה בעל עוצמה מקסימלית.