מבוא לסטטיסטיקה ותכנון ניסויים

פרק 13 - התפלגות הדגימה ומשפט הגבול המרכזי

תוכן העניינים

1. התפלגות ממוצע המדגם ומשפט הגבול המרכזי 1
2. התפלגות סכום תצפיות בלתי תלויות ומשפט הגבול המרכזי 8

התפלגות ממוצע המדגם ומשפט הגבול המרכזי:

רקע:

בפרק זה נדון בהתפלגות של ממוצע המדגם: $\bar{x}=\frac{\sum x_i}{n}$.

מכיוון שממדגם למדגם אנו יכולים לקבל ממוצע מדגם שונה, אזי ממוצע המדגם הוא משתנה מקרי ויש לו התפלגות.

גדלים המתארים התפלגות כלשהי או אוכלוסייה כלשהי נקראים פרמטרים.

להלן רשימה של פרמטרים החשובים לפרק זה:

ממוצע האוכלוסייה נסמן ב- μ (נקרא גם תוחלת).

שונות אוכלוסייה נסמן ב- σ^2.

סטיית תקן של אוכלוסייה: σ.

תכונות התפלגות:

ממוצע כל ממוצעי המדגם האפשריים שווה לממוצע האוכלוסייה: $E(\bar{x})=\mu_{\bar{x}}=\mu$.

שונות כל ממוצעי המדגם האפשריים שווה לשונות האוכלוסייה מחולק ב- n.

תכונה זו נכונה רק במדגם מקרי: $V(\bar{x})=\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}$.

יש יחס הפוך בין גודל המדגם לבין שונות ממוצעי המדגם.

אם נוציא שורש לשונות נקבל סטיית תקן של ממוצע המדגם שנקראת גם

טעות תקן: $\sigma(\bar{x})=\sqrt{\frac{\sigma^2}{n}}=\frac{\sigma}{\sqrt{n}}$.

דוגמה (פתרון בהקלטה):

השכר הממוצע במשק הינו 9000 ₪ עם סטיית תקן של 4000. דגמו באקראי 25 עובדים.

א. מיהי אוכלוסיית המחקר? מהו המשתנה הנחקר?
ב. מהם הפרמטרים של האוכלוסייה?
ג. מה התוחלת ומהי סטית התקן של ממוצע המדגם?

© כל הזכויות שמורות לברק קנדל

דגימה מהתפלגות נורמאלית:

אם נדגום מתוך אוכלוסייה שהמשתנה בה מתפלג נורמאלית עם ממוצע μ ושונות σ^2.

ממוצע המדגם גם יתפלג נורמאלית: $Z_{\bar{x}} = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$, $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

דוגמה (פתרון בהקלטה) :

משקל תינוק ביום היוולדו מתפלג נורמאלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.
מה ההסתברות שבמדגם של 4 תינוקות אקראיים בעת הולדתם המשקל הממוצע של התינוקות יהיה מתחת ל-3.5 ק"ג?

משפט הגבול המרכזי:

אם אוכלוסייה מתפלגת כלשהו עם ממוצע μ ושונות σ^2 אזי עבור מדגם מספיק גדול $(n \geq 30)$ ממוצע המדגם מתפלג בקירוב נורמאלי: $\bar{x} \sim\succ N\left(\mu, \frac{\sigma^2}{n}\right)$.

דוגמה (פתרון בהקלטה):

משקל חפיסת שוקולד בקו ייצור מתפלג עם ממוצע 100 גרם וסטיית תקן של 4 גרם.
דגמו מקו הייצור 36 חפיסות שוקולד אקראיות.
מה ההסתברות שהמשקל הממוצע של חפיסות השוקולד שנדגמו יהיה מתחת ל-102 גרם?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.

א. מיהי האוכלוסייה?

ב. מה המשתנה?

ג. מהם הפרמטרים?

ד. מהו גודל המדגם?

ה. מהו תוחלת ממוצע המדגם?

ו. מהי טעות התקן?

2) להלן התפלגות מספר מקלטי הטלוויזיה למשפחה בישוב מסוים:

מספר המשפחות	מספר מקלטים
500	0
2500	1
3500	2
3000	3
500	4
סך הכול $N=10000$	

א. בנו את פונקציית ההסתברות של X.

ב. חשבו את התוחלת, השונות וסטיית התקן של X.

ג. אם נדגום 4 משפחות מהישוב עם החזרה מה תהיה התוחלת, מהי השונות ומהי סטיית התקן של ממוצע המדגם?

3) אם נטיל קובייה פעמיים ונתבונן בממוצע התוצאות שיתקבלו, מה תהיה התוחלת ומה תהיה סטיית התקן של ממוצע זה?

4) משקל תינוק ביום היוולדו מתפלג נורמאלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.

א. מה ההסתברות שתינוק אקראי בעת הלידה ישקול פחות מ-3800 גרם?

נתון כי ביום מסוים נולדו 4 תינוקות.

ב. מה ההסתברות שהמשקל הממוצע שלהם יעלה על 4 ק"ג?

ג. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה מתחת ל-2.5 ק"ג?

ד. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה רחוק מהתוחלת בלא יותר מ-50 גרם?

ה. הסבירו ללא חישוב כיצד התשובה לסעיף הקודם הייתה משתנה אם היה מדובר על יותר מ-4 תינוקות?

© כל הזכויות שמורות לברק קנדל

5) הגובה של המתגייסים לצה"ל מתפלג נורמאלית עם תוחלת של 175 ס"מ וסטיית תקן של 10 ס"מ. ביום מסוים התגייסו 16 חיילים.

א. מה ההסתברות שהגובה הממוצע שלהם יהיה לפחות 190 ס"מ?

ב. מה ההסתברות שהגובה הממוצע שלהם יהיה בדיוק 180 ס"מ?

ג. מה ההסתברות שהגובה הממוצע שלהם יסטה מתחולת הגבהים בפחות מ-5 ס"מ?

ד. מהו הגובה שבהסתברות של 90% הגובה הממוצע של המדגם יהיה נמוך ממנו?

6) הזמן הממוצע שלוקח לאדם להגיע לעבודתו 30 דקות עם שונות של 16 דקות רבועות. האדם נוסע לעבודה במשך שבוע 5 פעמים.
לצורך הפתרון הניחו שזמן הנסיעה לעבודה מתפלג נורמאלית.

א. מה ההסתברות שבמשך שבוע משך הנסיעה הממוצע יהיה מעל 33 דקות?

ב. מהו הזמן שבהסתברות של 90% ממוצע משך הנסיעה השבועי יהיה גבוה ממנו?

ג. מה ההסתברות שממוצע משך הנסיעה השבועי יהיה מרוחק מ-30 דקות בלפחות 2 דקות?

ד. כיצד התשובה לסעיף הקודם הייתה משתנה אם האדם היה נוסע לעבודה 6 פעמים בשבוע?

7) נפח היין בבקבוק מתפלג נורמאלית עם תוחלת של 750 סמ"ק וסטיית תקן של 10 סמ"ק.

א. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה בדיוק 755 סמ"ק?

ב. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה יותר מ-755 סמ"ק?

ג. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה לפחות 755 סמ"ק?

ד. בקבוקיי היין שבארגז נמזגים לקערה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקערה?

8) משתנה מתפלג נורמאלית עם תוחלת 80 וסטיית תקן 4.

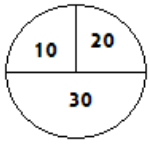
א. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה כאשר גודל המדגם הוא 9?

ב. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה שגודל המדגם הוא 16?

ג. הסבר את ההבדל בתשובות של שני הסעיפים.

© כל הזכויות שמורות לברק קנדל

9) בקזינו ישנה רולטה. על הרולטה רשומים המס׳ הבאים כמוראה בשרטוט:



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.

א. בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד.

ב. מה התוחלת ומה השונות של סכום הזכייה?

ג. אם האדם ישחק את המשחק 5 פעמים מה התוחלת ומה השונות של ממוצע סכום הזכייה בחמשת המשחקים?

ד. אם האדם משחק את המשחק 50 פעם מה ההסתברות שבסה״כ יזכה ב-1050 ₪ ומעלה?

10) לפי הערכות הלשכה המרכזית לסטטיסטיקה השכר הממוצע במשק הוא 8000 ₪ עם סטיית תקן של 3000 ₪.
מה ההסתברות שבמדגם מקרי של 100 עובדים השכר הממוצע יהיה יותר מ-8500 ₪?

11) קובייה הוטלה 50 פעמים.
מה ההסתברות שהממוצע של התוצאות יהיה לפחות 3.72?

12) אורך צינור שמפעל מייצר הינו עם ממוצע של 70 ס״מ וסטיית תקן של 10 ס״מ.

א. נלקחו באקראי 100 מוטות, מה ההסתברות שממוצע אורך המוטות יהיה בין 68 ל-78 ס״מ?

ב. יש לחבר 2 בניינים באמצעות מוטות. המרחק בין שני הבניינים הינו 7200 ס״מ. מה ההסתברות ש-100 המוטות יספיקו למלאכה?

ג. מה צריך להיות גודל המדגם המינימאלי, כדי שבהסתברות של 5% ממוצע המדגם יהיה קטן מ-69 ס״מ. היעזרו במשפט הגבול המרכזי.

13) נתון משתנה מקרי בדיד בעל פונקציית ההסתברות הבאה:

X	8	6	4	2
$P(X)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

מתוך התפלגות זו נלקח מדגם מקרי בגודל 50.
מה הסיכוי שממוצע המדגם יהיה קטן מ-5?

© כל הזכויות שמורות לברק קנדל

14) נתון ש- $X \sim N\left(\mu,\sigma^2\right)$. דגמו 5 תצפיות מאותה התפלגות והתבוננו בממוצע המדגם $\bar{X}$ לכן: $P\left(\bar{X} > \mu\right)$ יהיה (בחרו בתשובה הנכונה):

א. 0.
ב. 0.5.
ג. 1.
ד. לא ניתן לדעת.

15) נתון ש- X מתפלג כלשהו עם תוחלת μ ושונות σ^2.
החליטו לבצע מדגם בגודל 200 מתוך ההפלגות הנתונה לפי משפט הגבול המרכזי מתקיים (בחרו בתשובה הנכונה):

א. $X \sim N\left(\mu,\frac{\sigma^2}{200}\right)$.
ב. $\mu \sim N\left(\mu,\frac{\sigma^2}{200}\right)$.
ג. $\bar{X} \sim N\left(\mu,\sigma^2\right)$.
ד. $\bar{X} \sim N\left(\mu,\frac{\sigma^2}{200}\right)$.

16) נתון ש- $X \sim N\left(\mu,\sigma^2\right)$. אם נדגום n תצפיות מתוך ההתפלגות ונגדיר: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$,
אזי (בחרו בתשובה הנכונה):

א. μ ו- $\bar{X}$ יהיו משתנים מקריים.
ב. μ יהיה משתנה מקרי ו- $\bar{X}$ קבוע.
ג. $\bar{X}$ יהיה משתנה מקרי ו- μ קבוע.
ד. μ ו- $\bar{X}$ יהיו קבועים.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. כלל הסטודנטים במכללה שסיימו סטטיסטיקה א. ב. ציון.
ג. ממוצע: 78, סטיית תקן: 15. ד. 2.
ה. 78. ו. 10.6.

2) א. להלן טבלה:

X	0	1	2	3	4
$P(X)$	0.05	0.25	0.35	0.3	0.05

ב. $\mu = 2.05$, $\sigma^2 = 0.9475$, $\sigma = 0.973$

ג. $\mu_{\bar{x}} = 2.05$, $\sigma^2_{\bar{x}} = 0.2369$, $\sigma(\overline{X}) = 0.486$.

3) $\mu_{\bar{x}} = 3.5$, $\sigma(\overline{X}) = 1.21$.

4) א. 0.8413. ב. 0.0013. ג. 0. ד. 0.1974.

5) א. 0. ב. 0. ג. 0.9544. ד. 178.205.

6) א. 0.0465. ב. 27.71. ג. 0.2628. ד. התשובה הייתה קטנה.

7) א. 0. ב. 0.1587. ג. 0.1587. ד. 0.5.

8) א. 0.5468. ב. 0.6826.

9) א. להלן טבלה:

X	10	20	30
$P(X)$	0.25	0.25	0.5

ב. התוחלת: 22.5, השונות: 68.75. ג. התוחלת: 22.5, השונות: 13.75.
ד. 0.8997.

10) 0.0475.

11) 0.1814.

12) א. 0.9772. ב. 0.0228. ג. 271.

13) 0.5.

14) ב'.

15) ד'.

16) ג'.

© כל הזכויות שמורות לברק קנדל

התפלגות סכום תצפיות בלתי תלויות ומשפט הגבול המרכזי:

רקע:

כעת נדון בסטטיסטי המבטא את סכום התצפיות במדגם: $T=\sum_{i=1}^{n}X_i$.

כאשר כל התצפיות נדגמו באקראי מאותה אוכלוסייה, כלומר, היו: $X_1,\ldots,X_n$ -

משתנים מקריים בלתי תלויים בעלי התפלגות זהה שתוחלתה μ ושונותה σ^2 אזי:

התוחלת והשונות של סכום התצפיות: $E(T)=n\mu$, $V(T)=n\sigma^2$.

דגימה מתוך התפלגות נורמלית:

אם: $X\sim N(\mu,\sigma^2)$, אזי: $Z=\frac{T-n\mu}{\sqrt{n\sigma^2}}$, $T\sim N(n\mu,n\sigma^2)$.

משפט הגבול המרכזי:

אם X מתפלג כלשהו וידוע כי: $V(X)=\sigma^2$, $E(X)=\mu$,

אזי עבור מדגם מספיק גדול (לפחות 30): $T\sim\succ N(n\mu,n\sigma^2)$.

דוגמה (פתרון בהקלטה):

בעיר מסוימת המשכורת הממוצעת של עובד הינה 8000 ₪. עם סטיית תקן של 2000 ₪.
נדגמו 100 עובדים מהעיר שמפקידים את משכורותיהם לסניף בנק.

1. מה התוחלת וסטיית התקן של סך המשכורות שיופקדו לסניף הבנק על ידי העובדים הללו?
2. מה ההסתברות שלסניף יופקד פחות מ-780 אלף ₪ ע"י אותם עובדים?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) המשקל באוכלוסייה מסוימת מתפלג נורמאלית עם תוחלת של 60 ק"ג וסטיית תקן של 10 ק"ג.

א. מה הסיכוי שאדם אקראי מהאוכלוסייה ישקול מתחת ל-65 ק"ג?

ב. מה הסיכוי שהמשקל הממוצע של 4 אנשים אקראיים יהיה מתחת ל-65 ק"ג?

ג. מה הסיכוי שהמשקל הכולל של 4 אנשים אקראיים יהיה מתחת ל-240 ק"ג?

2) נפח יין בבקבוק מתפלג נורמאלית עם תוחלת של 750 מ"ל וסטיית תקן של 20 מ"ל. אדם קנה מארז של 4 בקבוקי יין.

א. מהי התוחלת ומהי סטיית התקן של נפח היין במארז?

ב. את היין שבמארז האדם מזג לכלי שקיבולתו 3.1 ליטר.
מה ההסתברות שהיין יגלוש מהכלי?

ג. אם לא היה נתון שנפח היין מתפלג נורמאלית. האם התשובה לסעיף א' הייתה משתנה? האם התשובה לסעיף ב' הייתה משתנה?

3) בספר כלשהו 500 עמודים. קצב הקריאה הממוצע הוא עמוד אחד ב-4 דקות עם סטיית תקן של 1 דקות.

א. מה ההסתברות לסיים את הפרק הראשון (40 עמודים) תוך שעתיים וחצי?

ב. מהו האחוזון ה-95 לזמן סיום קריאת הספר?

תשובות סופיות:

1) א. 0.6915. ב. 0.8413. ג. 0.5 .

2) א. תוחלת 3000 מ"ל, סטיית תקן 40 מ"ל. ב. 0.0062.

ג. סעיף א'- לא משתנה, סעיף ב'- לא פתיר, התבסס על התפלגות נורמלית.

3) א. 0.0571. ב. 2036.8.

© כל הזכויות שמורות לברק קנדל