

הסברים לתשובות למבחן לדוגמא 1- יוני 2016

1. על מנת שסטטיסטי יהווה אומדן בלתי מוטה לפרמטר, התוחלת שלו צריכה להיות שווה לפרמטר. הממוצע הינו אומדן בלתי מוטה שכן אם נדגום אינסוף מדגמים בגודל מסוים ונחשב עבור כל מדגם את הממוצע שלו, הממוצע של כל הממוצעים הנ"ל יהיה שווה לממוצע הכללי של אותה האוכלוסייה ממנה נדגמו המדגמים, כלומר לתוחלת האוכלוסייה אותה הממוצע מבקש לאמוד. השונות לעומת זאת היא אומדן מוטה כלפי מטה שכן אם נבצע את אותה הפעולה ונחשב עבור כל מדגם את השונות שלו, ממוצע כל השונות יהיה קטן יותר מהשונות הכללית של האוכלוסייה. טווח המדגם גם הוא אומדן מוטה שכן הוא מושפע אך ורק מהתוצאות הקיצוניות שנדגמו בכל מדגם, על כן התשובה הנכונה היא א' בלבד.

2. על מנת לדעת כיצד הוצאת או הוספת תצפיות משפיעה על סטיית התקן/השונות עלינו לבדוק האם התצפיות רחוקות מהממוצע ביותר או פחות מסטיית תקן אחת. במקרה המובא בשאלה ציוני התקן של התצפיות שהוצאו נחשבות קרובות לממוצע היות והן רחוקות מהממוצע בפחות מסטיית תקן. היות ותצפיות הקרובות לממוצע מקטינות את השונות, הוצאה של תצפיות שכאלו מגדילה את השונות ולכן השונות הגולמית תגדל לאחר הוצאת ציוניהם של ויקטור וסרגיי. היות וסטיית התקן גדלה, וציונו הגולמי של המפקד שמגר נשמר, בחישוב ציון התקן שלו כעת המונה נותר כמו שהוא (מרחק התצפית מהממוצע) אך המכנה גדל (סטיית התקן) ולכן ציון התקן של המפקד שמגר הפך להיות קטן יותר. בהתפלגות ציוני תקן הממוצע תמיד שווה לאפס והשונות וסטיית התקן שוות לאחת ולכן סטיית התקן של ציוני התקן לא תשתנה ולכן התשובה הנכונה היא ד'.

3. השאלה למעשה מתייחסת לעוצמת המבחן ולסיכוי לטעות מסוג ראשון. הסיכוי לדחות את השערת האפס, במידה והיא אינה נכונה מוגדר כעוצמת המבחן. עוצמת המבחן תלויה במספר גורמים, ביניהם דרגות החופש. במקרה זה ככל שמספר דרגות החופש גדל, כך הערך הקריטי קטן ועוצמת המבחן גדלה. יותר דרגות חופש = יותר טוב! לכן תשובה ב' נכונה. הסיכוי לטעות מסוג ראשון לעומת זאת (אשר שווה לאלפא) אינו מושפע משום דבר למעט החלטת החוקר. ולכן תשובות ג' וד' אינן נכונות.

4. מתאם ספירמן בודק קשר בין שני משתנים מסולם סדר ומעלה. היות ומתאם ספירמן בודק אך ורק את הקשר בין דירוגי התצפיות בשני המשתנים ללא כל התייחסות לרווחים ביניהם, איננו יכולים ללמוד מעוצמת הקשר במתאם ספירמן על עוצמת הקשר כפי שיחושב במתאם פירסון ועל כן התשובה היא ד. במידה והמצב היה הפוך והיינו יודעים שמתאם פירסון הוא מושלם היינו יכולים לדעת בוודאות שגם מתאם ספירמן יהיה מושלם.

5. במדגם בודד יכולה להתקבל כל סטיית תקן היות ומדובר על מדגם מקרי אשר אינו בהכרח מייצג את כלל האוכלוסייה ואינו בהכרח מכיל את מגוון התצפיות השונות הקיימות באוכלוסייה ולכן איננו יכולים לדעת מה יהיה היחס בין סטיית התקן שהתקבלה במדגם לסטיית התקן באוכלוסייה. הדבר היחידי שאנחנו יודעים הוא שברמת התוחלת (כלומר במידה והיינו דוגמים אינסוף מדגמים באותו הגודל ומחשבים את הממוצע של כל סטיות התקן שהתקבלו) סטיית התקן של המדגם תהיה קטנה יותר מסטיית התקן של האוכלוסייה. היות והשאלה שואלת על רמת המדגם הבודד ולא על רמת התוחלת התשובה היא ד.

6. היות וסמי וסוסו שניהם מבצעים בדיקת השערות באותה רמת בטחון, הרי שלשניהם אלפא זהה. נתון בשאלה שידוע כי השערת האפס נכונה ולכן הסיכוי לדחות את השערת האפס (בטעות כמובן) היות ונתון כי היא נכונה) שווה לאותה האלפא ולכן לשניהם סיכוי שווה לדחות את השערת האפס והתשובה הנכונה היא ג. בשאלות מסוג זה יש לשים לב האם נתון שהשערת האפס נכונה או לא ובהתאם לבחור בהסתברויות המתאימות לדחיית/אי-דחיית השערת האפס.

7. על מנת לבחון באמצעות איזה משתנה כדאי לנבא את הרווח עלינו לבדוק לאיזה משתנה יש מתאם פירסון חזק יותר עם הרווח. היות וההפרשים בין התצפיות השונות בתוך משתנה הותק בעבודה שווים להפרשים בין התצפיות התואמות בתוך משתנה שטח החדר, הרי שהמתאם שיחושב יהיה זהה עבור שני המשתנים ולכן הניבוי באמצעות שני המשתנים יהיה זהה והתשובה היא ג.

8. עלינו לחשב רווח סמך בהתאם לנתוני השאלה ולנוסחת רווח סמך עבור Z המתאים לרמת בטחון של 95%. הרווח המתקבל מתאים לתשובה ב.

9. בהתאם לתשובה לשאלה 3, במידה והשערת האפס אינה נכונה הסיכוי לדחות אותה ולקבל תוצאה מובהקת שווה לסיכוי לטעות מסוג ראשון ולכן יהיה זהה בשני המקרים והתשובה הנכונה היא ג.

10. הנבדקים שקיבלו ציון 72 היו רחוקים מהממוצע ב-8 נקודות כלומר ביותר מסטיית תקן אחת ולכן תרמו להגדלת השונות. היות ולאחר התוספת של 5 נקודות ציונם שווה ל-77 הם כעת קרובים יותר לממוצע ולכן השונות תקטן. ניתן לראות את השפעת התוספת גם באמצעות מבט על הנוסחה לחישוב השונות- היות והמכנה נותר זהה והמונה קטן השונות כולה תקטן ולכן התשובה היא ב.

11. נחפש מחקר העומד בהנחות מבחן t למדגמים תלויים: שני מדגמים עם תצפיות תלויות או מדגם אחד של זוגות של תצפיות כאשר התפלגות הדגימה של הממוצעים מתפלגת נורמלית והמשתנה התלוי הוא בסולם רווח או מנה. במסיח א מתואר מחקר עם 4 רמות ולכן הוא לא מתאים. במסיח ב המשתנה הוא אמנם בסולם מנה, אך בגלל שהוא מתפלג א-סימטרית באוכלוסייה והוא נמדד רק על 20 זוגות של תצפיות, התפלגות הדגימה לא תהיה נורמלית ולכן תשובה זו אינה נכונה. במסיח ג המשתנה התלוי נמדד בסולם מנה וישנם 42 זוגות של תצפיות על כן התפלגות הדגימה תתפלג נורמלית. כמו כן מדובר על זוגות שיש ביניהם קשר ולכן יש לבצע מבחן T למדגמים תלויים ולכן זוהי התשובה הנכונה. במסיח ד המשתנה התלוי נמדד בסולם סדר ולכן אינו מתאים כלל לביצוע מבחן T.

12. היות ומספר הבורקסים שאוכלים חיילים מתפלג נורמלית, ביכולתנו להשתמש בטבלת Z כדי לעבור מציוני תקן לאחוזונים וכך לענות על השאלה. כדי לדעת מהו מספר הבורקסים ש-15% מהחיילים אוכלים יותר ממנו, עלינו לבדוק ראשית מהו ציון התקן שמעליו ישנם 15% מההתפלגות. ממבט בטבלת Z מספר זה שווה בקירוב ל-1. כעת עלינו לבדוק מהו הציון הגולמי בהתפלגות הנתונה לנו, המתאים לציון תקן של 1. נעשה זאת על ידי שימוש בנוסחה לחישוב ציון

$$\text{תקן: } Z_x = \frac{x_i - \bar{x}}{S_x} \text{ נציב: } 1 = \frac{x_i - 32}{7} \text{ ומכאן שהציון אותו אנו מחפשים שווה ל-39, כלומר}$$

תשובה ג היא הנכונה.

13. הביטוי "פספוס מציאת אפקט אמיתי" הוא שם אחר ל"טעות מסוג שני" כלומר לבטא. נבדוק כיצד כל פעולה משפיעה על בטא. הגדלת אלפא גורמת להקטנת בטא ולכן זוהי התשובה הנכונה. הקטנת המדגם גורמת להגדלת שונות התפלגות הדגימה, כלומר להגדלת הרעש בניסוי ולכן גורמת להגדלת בטא. ניסוח השערה זו צדדית מגדיל את הסיכוי שנפספס אפקט היות ואנו מחלקים את אלפא לשני צידי ההתפלגות ולכן גם הוא מגדיל את בטא. ביצוע מבחן t במקום מבחן z מגדיל את הערך הקריטי ולכן מגדיל את הסיכוי לפספס אפקט, כלומר את בטא.

14. ראשית על מנת לזהות משתנה בלתי תלוי עלינו לחפש גורם שלו יש לפחות שתי רמות בניסוי והוא יכול להיות נמדד או מתופעל אך ברצוננו לבדוק כיצד הוא משפיע על משתנה אחר. משתנה תלוי

יהיה תמיד משתנה אשר השתנותו נמדדת כתלות במשתנה הבלתי תלוי. בניסוי נבדקים שני סוגי שוקולד שונים ולכן מדובר על משתנה. סוג השוקולד הוא משתנה שאינו מכיל מידע למעט שמו ולכן מדובר על משתנה מסולם שמי. כנייל לגבי מקום הנסיעה, אשר גם הוא משתנה שמי. כמות הילדים ומהירות ההתכה מהווים שניהם משתנים תלויים. שניהם מסולם מנה היות ויש משמעות לאפס בשניהם, כלומר אפס ילדים הכוונה להיעדר ילדים באופן מוחלט, כפי שאפס מהירות משמעותה היעדר מהירות לחלוטין. לכן התשובה הנכונה היא ג.

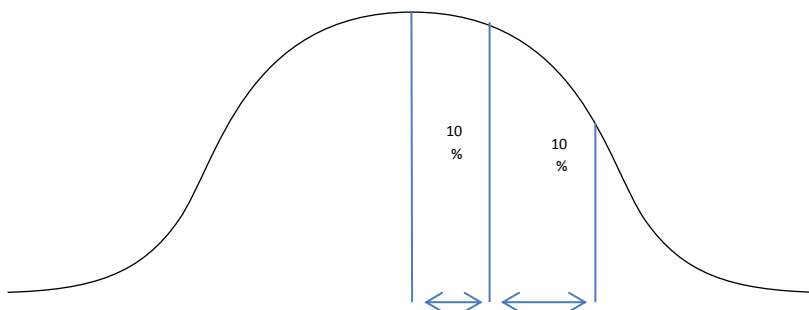
15. בהתפלגות אסימטרית חיובית המדדים מתפרשים לפי הסדר: אמצע טווח < ממוצע < חציון < שכיה (שחמ"ט) ולכן על הממוצע להיות גדול מהשכיה ומהחציון כלומר גדול מ87. התשובה היחידה המתאימה ד.

16. כדי שאורייך יחשב טוב באופן יחסי בהטלת כדור ספוג כשם שהוא בריצת 60 מטר, על ציון התקן שלו בשני המקצועות להיות זהה, כלומר 1.5. כל שנותר לנו לעשות הוא לחשב מהו ציון הגלם המתאים לציון תקן של 1.5 עבור הנתונים של הטלת כדור ספוג, כלומר עבור ממוצע 24 וס"ת 2. נחשב לפי הנוסחה: $Z_x = \frac{x_i - \bar{x}}{S_x}$. עם הנתונים: $1.5 = \frac{x_i - 24}{2}$ ולכן התשובה הנכונה היא 27, כלומר ג.

17. סוג המכונית של אדם והעדפתו הפוליטית הם שניהם משתנים שמיים ולכן המבחן היכול לבחון את הקשר ביניהם הוא רק מבחן א-פרמטרי, כלומר מבחן חי בריבוע כפי שמופיע בתשובה ב.

18. היות ולא נתון בשאלה ששונות ציוני המתמטיקה ידועה אין באפשרותנו להשתמש במבחן Z אלא רק במבחן t. היות ואותם תלמידים נמדדו לפני ואחרי השיעורים הרי שעלינו להשתמש במבחן t למדגמים תלויים ולכן התשובה הנכונה היא ב.

19. להתפלגות נורמלית יש צורה של פעמון. לכן ככל שנתרחק מהמרכז נראה שגובה ההתפלגות נמוך יותר. השטח מתחת להתפלגות שווה לשכיחות או במילים אחרות לאחוז התצפיות מההתפלגות הנמצאות בטווח ערכים מסוים. לכן ככל שנתרחק ממרכז ההתפלגות, כך נקבל עבור אותו השטח גובה נמוך יותר וכדי לפצות על הירידה בגובה ולשמור על אותו השטח נצטרך להגדיל את המרחק על ציר ה-X, כלומר את המרחק. מסיבה זו המרחק בין האחוזון ה-60 לאחוזון ה-70 (אשר נמצא רחוק יותר מהמרכז ולכן גובה ההתפלגות בשטח זה קטן יותר) גדול יותר מהמרחק בין האחוזון ה-50 לאחוזון ה-60 (הנמצא קרוב יותר למרכז ולכן גובה ההתפלגות בשטח זה גדול יותר).



20. הנוסחא לסטיית התקן של התפלגות הדגימה היא $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$ ולכן אם גודל המדגם גדול מ-1 הרי

שסטיית התקן של התפלגות הדגימה תהיה קטנה יותר מסטיית התקן של האוכלוסייה ולכן התשובה הנכונה היא ג.

21. כאמור, ציונים הקרובים לממוצע תורמים להקטנת השונות בעוד ציונים הרחוקים מהממוצע תורמים להגדלת השונות. היות והוסרו שני ציונים הזהים לממוצע הרי שהם תרמו להקטנת השונות ולכן הסרתם תהפוך את השונות לגדולה יותר ולכן תשובה ב נכונה.

22. גובה מעל פני הים הוא משתנה בסולם רווח היות ויש משמעות להפרשים בין התצפיות אך אין משמעות לאפס המוחלט (היות והוא נמדד ביחס לפני הים ולא מציין העדר גובה לחלוטין). התוספת של 4 מטרים נעשתה בצורה אחידה לכל המדידות ולכן היא מהווה טרנספורמציה ליניארית של הוספת קבוע. טרנספורמציה ליניארית של הוספת קבוע אינה משנה את תכונות המשתנה וגם לא את סולם המדידה ולכן המשתנה היה ונותר בסולם רווח ותשובה ב היא הנכונה.

23. על מנת לחשב את אחוז ההתפלגות אשר נמצא מתחת לציון תקן אחת ומטה מהממוצע עלינו לבדוק בטבלת Z מהו אחוז התצפיות הנמצאות מתחת לציון תקן 1-. במידה ובטבלה הנתונה לנו לא מופיעים ציוני תקן שליליים אלא רק ציוני תקן חיוביים נשתמש בתכונת הסימטריות של ההתפלגות הנורמלית ובמקום לבדוק מה אחוז התצפיות הנמצאות מתחת לציון תקן 1- נבדוק מהו אחוז התצפיות הנמצאות מעל ציון תקן 1. מבט בטבלה יראה כי מעל ציון תקן 1, ישנם 15.87% מהתצפיות ולכן זהו המספר שעלינו לחפש בתשובות. לאחר עיגול התשובה המתקבלת היא 16% כלומר תשובה ג.

24. המעבר מס"מ למטרים משמעותו חלוקת כל הערכים ב-100. פעולה זו מהווה טרנספורמציה ליניארית של חלוקה בקבוע שהיא כאמור לא משנה את מהות התופעה וגם אינה משנה את הסיכוי לדחות את השערת האפס. הסיבה לכך היא שלמרות שאכן החלוקה תקטין את השונות במחקר, היא תקטין גם את הרווחים בין התצפיות באותה הפרופורציה ולכן פעולה זו לא תשפיע כלל וחוקר א' טועה. לעומת זאת הגדלת המדגם כידוע מגדילה את עוצמת המבחן היות והיא מקטינה את שונות התפלגות הדגימה ולכן חוקר ב' צודק ותשובה ב נכונה.

25. היות והמשתנה הנמדד (מהירות הריצה של ילדים) הוא משתנה מסולם מנה, המבחן המתאים לבדיקת השערתו של החוקר הוא מבחן פרמטרי כלומר מבחן t או מבחן z. היות ולא נתון לנו כי שונות האוכלוסייה ידועה לא נוכל להשתמש במבחן z אלא רק במבחן t והיות והחוקר דגם שני מדגמים שונים ובלתי תלויים של 50 ילדים בכל קבוצה על החוקר לבצע מבחן t למדגמים בלתי תלויים כלומר תשובה ג היא התשובה הנכונה.