**דף נוסחאות מורחב**

**סטטיסטיקה תיאורית**

סימנים בסיסיים לשימוש :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| אורך האינטרוול | *l* |  | אמצע הקבוצה | *or MP* |
| צפיפות האינטרוול | *d* | גודל המדגם | *N* |
| ממוצע |  | סכום אברים |  |
| חציון |  | שכיחות מוחלטת | *f* |
| השכיח |  | שכיחות מצטברת | *F* |
| סטיית תקן |  | שכיחות יחסית | *p* |
| שונות |  | ציון תקן |  |

n - מספר קבוצות (מחלקות)

Li - גבול תחתון של קטע i.

Ri - גבול עליון של קטע i.



מציאת טווח:

1) 

(קיבוץ תצפיות לפי ערכים בודדים)

2) 

(קיבוץ נתונים לפי קטעים.)





 

מציאת שכיח ( קיבוץ תצפיות לפי קטעים)



 - קבוצת השכיח 

מציאת חציון (קיבוץ תצפיות לפי קטעים)



 - קבוצת החציון ( ו-  ).

אם ניתן למצוא את i שעבורו 



מציאת חציון (קיבוץ תצפיות לפי ערכים בודדים)

אם ניתן למצוא את i שעבורו  ו-  .



אם ניתן למצוא את i שעבורו 





**אקסיומות ההסתברות**

1) תכונות כלליות של פונקצית הסתברות **** כאשר - מאורע כלשהו

 - מרחב מאורעות או אוסף כל התוצאות האפשריות של ניסוי או מאורע וודאי.  
 ****2) ,  כאשר  - מאורע משלים 

3)   
4)  מאורעות זרים אם  ומתקיים 

**נוסחאות שימושיות נוספות**:

5)    
6) ,  (דה מורגן)

**שיטת ריבוע הקסם:**

מציירים ריבוע. בקצה השמאלי כותבים  ובקצה העליון כותבים , כך נקבל שתי שורות ושתי עמודות.

בקצה הימני של השורה של כותבים ,

בקצה הימני של השורה של כותבים ,

בקצה התחתון של העמודה שלכותבים ,

בקצה התחתון של העמודה של כותבים ,

בפינה הימנית התחתונה כותבים 1,

בארבע המשבצות באמצע כותבים את הערכים:

 בשורה של  ובעמודה של 

 בשורה של  ובעמודה של 

 בשורה של  ובעמודה של 

 בשורה של  ובעמודה של 

נשים לב שסכום שני המספרים בצד שמאל בכל שורה שווה למספר המופיע בקצה הימני של השורה. סכום שני המספרים למעלה בכל עמודה שווה למספר המופיע בקצה התחתון של כל עמודה.

אם נתונים שלושה ערכים בריבוע, כך שלא כל השלושה באותה שורה או באותה העמודה, ניתן לחשב את שאר הערכים בעזרתם.

**קומבינטוריקה**

תמורות

****

בחירות

|  |  |  |
| --- | --- | --- |
| **בלי החזרה** | **עם החזרה** |  |
|  |  | **סדר חשוב**  (חליפות) |
|  |  | **סדר לא חשוב**  (צירופים) |

חישוב של מספר אפשרויות למאורעות שלא יקרו נעשה על ידי הפחתת מספר האפשרויות שהמאורעות כן יקרו מכלל האפשרויות (למשל מספר האפשרויות ששני אנשים לא ישבו ביחד מחשבים על ידי הפחתה מכלל האפשרויות את מספר האפשרויות ששני האנשים כן ישבו ביחד).

חישוב של הסתברות לפעולה מסוימת מתוך כלל הפעולות מחושב על ידי חישוב מספר האפשרויות לפעולה המסוימת חלקי מספר האפשרויות עבור כלל הפעולות.

**הסתברות מותנית**

- קוראים: ההסתברות המותנה של מאורע כאשר ידוע שהמאורע 

התרחש (בקיצור, ההסתברות של  בתנאי ).

, 

אם  מאורעות זרים אז מתקיים 

אם  מאורעות בלתי תלויים אז מתקיים  , 

אם  מאורעות תלויים אז מתקיים , 

**נוסחת ההסתברות השלמה:**



כאשר  מאורעות זרים שאיחודם הוא כל המרחב : 

**נוסחת בייז :**



**נוסחה נוספת**: 

**שיטת העצים:**

בונים עץ כך שבהתחלה משרטטים n ענפים מסומנים ב-  שיוצאים מנקודה מסוימת, וכותבים על כל ענף את ההסתברויות המתאימות . אחר כך, מורידים מקצה של כל אחד מהענפים ענף נוסף המסומן ב- בהתאמה, וכותבים על כל אחד מהענפים החדשים את ההסתברויות . כעת, כדי לחשב P(A), מכפילים כל אחד מהערכים

, בערכים  שמעליו, ומחברים את הערכים מהענפים השונים שהתקבלו.

**מומנטים למשתנה בדיד**

תוחלת , שונות 

תכונות של תוחלת ושונות:



**התפלגות בינומית** (ברנולי) **

n - מספר ניסוים*,*  p - הסתברות של הצלחה בכל ניסוי בודד, נסמן: .

**ההסתברות** שמתוך n ניסויים יהיו בדיוק k הצלחות: 

**תוחלת ושונות** של התפלגות בינומית: 

"הצלחה בניסוי" היא הפעולה ש- X מודד, ללא קשר למשמעות שלו כדבר טוב (למשל אם רוצים לחשב הסתברות לבחור k מוצרים פגומים מתוך n מוצרים עם החזרה, אזי "הצלחה בניסוי" הוא לבחור מוצר פגום).

**התפלגות היפרגיאומטרית**

באוכלוסיה N פריטים, מתוכם D פריטים מיוחדים. מוציאים מהאוכלוסייה n פריטים

**ללא החזרה**. X = מספר פריטים מיוחדים שהוצאו.

ההסתברות שמתוך n פריטים יהיו k פריטים מיוחדים: .

תוחלת ושונות:



**התפלגות גיאומטרית** (עד הצלחה ראשונה) 

p - הסתברות של הצלחה בכל ניסוי בודד, נסמן: 

ההסתברות שההצלחה הראשונה תתרחש בניסוי מס' k: 

, , 

**סכום של סדרה הנדסית אינסופית**:  

**התפלגות פואסון**

התפלגות פואסון עם פרמטר : 

 הוא המספר הממוצע של האירועים ליחידת זמן, 

**ההסתברות** שביחידת הזמן יתרחשו בדיוק k אירועים היא: 

**תוחלת ושונות**: 

**קירוב פואסון לבינומית** (עבור n גדול ו p קטן) מציבים 

בהתפלגות פואסון חשוב להשתמש בממוצע ליחידת הזמן שרוצים לחשב, ואם בנתון המקורי של הממוצע הוא ליחדת זמן אחר, צריכים לחשב את הממוצע ליחידת הזמן שהתבקשנו בשאלה של חישוב ההסתברות (למשל, אם התבקשנו לחשב את ההסתברות שאירוע מסוים יתרחש מספר פעמים בשעה, ונתון שהאירוע מתרחש מספר פעמים בממוצע בדקה, צריך לחשב שיהיה מספר הפעמים שהאירוע יתרחש מספר הפעמים בממוצע בשעה).

**משתנה מקרי רציף**

(x) f פונקצית צפיפות, , (x) F פונקצית ההתפלגות.

, , , , ,

F(x) אינה יורדת, כלומר אם:  אז .  
 , 



**התפלגות נורמלית** (גאוס) 

X הוא משתנה נורמלי עם פרמטרים  ו-, אם פונקצית הצפיפות שלו ניתנת על ידי

סימונים: .

אם  אזי 

פונקצית התפלגות של X היא: 

כאשר  היא פונקצית ההתפלגות המצטברת של המשתנה הנורמלי הסטנדרטי

והיא נתונה בטבלה. תכונות של פונקצית :

, , .

, 

**משפט הגבול המרכזי**

יהיו- משתנים מקריים בלתי תלויים בעלי התפלגות זהה שתוחלתה  ושונותה  אזי

 או 

 או 

**קירוב נורמלי למשתנה בינומי**

יהי X משתנה בינומי בעל פרמטרים ו- כלומר , אזי עבור  מספיק גדול  מתקיים:



**סכום והפרש התפלגויות נורמליות**

יהיו ו- משתנים מקריים נורמליים בלתי תלויים כך ש: 



אזי מתקיים: 

****

**רווח בר- סמך לממוצע כאשר סטיית התקן ידועה:**

- ממוצע האוכלוסייה,  - שונות האוכלוסייה,  - סטיית התקן של האוכלוסייה.

**** - ממוצע המדגם,**** -רמת המובהקות,  רמת הביטחון, - גודל המדגם

רווח בר-סמך ל-  : 

**גודל המדגם** הדרוש כדי שברמת בטחון , הסטייה בין  לבין  לא תעלה על  הוא:



**רווח בר-סמך עבור פרופורציה**

רווח בר-סמךעבור פרופורציה**:** 

כאשר  היא פרופורציית ה"הצלחות" במדגם בגודל 

אם ידוע בוודאות ש - או אזי



גודל המדגם המבטיח שברמת סמך הסטייה בין לבין לא תעלה על



והוא: . כאשר הוא הערך הקרוב ביותר ל- 0.5 שיכול להתקבל.



אם לא שוללים את האופציה ש- קרוב ל- אזי מתקיים:



**אינטגרציה וגזירה בחלקים:**





