

אוניברסיטת אריאל בשומרון

החטיבה האקדמית - הפקולטה להנדסה

מבחן בקורס : הסתברות וסטטיסטיקה

מועד א' סמסטר א', שנה"ל תשפ"ב 3/2/2022

מספר הקורס 4422010-1,2,3

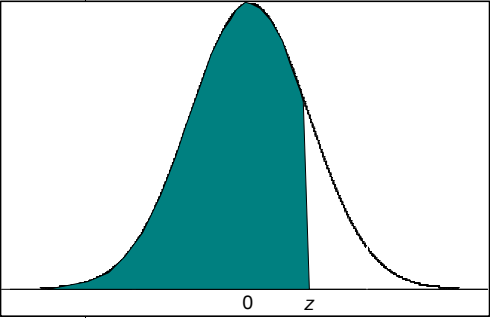
שם המרצה : ד"ר אליעזר הירשברג

מתרגלים: רן סגל, דביר רוס, צחי שביט

הוראות כלליות:

1. יש לענות על 5 מתוך 6 השאלות.
2. נא לסמן על כריכת המחברת איזה שאלות בחרת.
3. כל תשובה חייבת בנימוקים ברורים ומסודרים.
4. מחשבון כיס רגיל (לא גרפי).
5. משך הבחינה : 150 דקות.

-
1. במטבע א' ההסתברות לקבל עץ היא 0.6. במטבע ב' ההסתברות לקבל עץ היא 0.7. בוחרים באחד מהמטבעות באופן מקרי ומטילים אותו 10 פעמים.
א. מה ההסתברות לקבל בדיוק 7 פעמים עץ?
ב. אם התקבל 7 פעמים עץ, מה ההסתברות שנבחר מטבע א?
ג. חשב את התוחלת ושונות של $X =$ מספר הפעמים שהתקבל עץ ב-10 ההטלות.
 2. חבית שתייה במסיבה של 100 משתתפים מכילה 52 ליטר. כל משתתף במסיבה שותה בממוצע 0.5 ליטר עם סטיית תקן של 0.25 ליטר.
א. מה ההסתברות שהשתייה בחבית תספיק ל-100 המשתתפים?
ב. כמה ליטרים צריכה להכיל חבית השתייה כדי לספק את צרכי המשתתפים בהסתברות של 0.95?
 3. X הוא משתנה מקרי רציף עם פונקציית צפיפות $f(x) = A(4x - 2x^2)$, $0 < X < 2$.
א. חשב את הנעלם A
ב. חשב $P(X > 1)$
ג. מצא את פונקציית ההתפלגות $F(X)$.
 4. במפעל שתי מכוניות. ההסתברות שלפחות אחת מהן תפעל היא 0.6. אם הראשונה תפעל אז ההסתברות שהשנייה תפעל היא 0.5. אם השנייה תפעל אז ההסתברות שהראשונה תפעל היא 0.2.
א. חשב את ההסתברות ששתי המכוניות תפעל. (15 נק')
ב. מכניסים 3 כדורים ל-4 קופסאות באופן מקרי. מה ההסתברות שבדיוק קופסה אחת תהיה ריקה? (5 נק')
 5. בחנות מוכרים שני דגמים של מחשבוני. דגם יקר ב-60 ש"ח ודגם זול ב-40 ש"ח. ההסתברות שלקוח קונה את הדגם היקר היא 0.3 וההסתברות שלקוח קונה את הדגם הזול היא 0.7. ביום מסוים 20 לקוחות קנו מחשבוני. נגדיר $X =$ ההכנסות מ-20 לקוחות אלו. חשב את התוחלת והשונות של X .
 6. ראובן מציע לשמעון את המשחק הבא. אתה מטיל קובייה עד שתקבל תוצאה מספר 6. אני מטיל מטבע עד שאקבל עץ. אם מספר ההטלות שלך יהיה קטן ממספר ההטלות שלי אני אשלם לך 10 שקלים. אם מספר ההטלות שלי יהיה קטן ממספר ההטלות שלך אתה תשלם לי 2 שקלים. אם מספר ההטלות שלנו יהיה שווה אני אשלם לך שקל.
א. חשב את התוחלת של הרווח של שמעון במשחק.
ב. חשב את ההסתברות שהרווח של שמעון יהיה חיובי אחרי 30 משחקים.

Cumulative Probabilities for the Standard Normal (Z) Distribution										
						Values in the table correspond to the area under the curve of a standard normal random variable for a value at or below z .				
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998									
4.0	0.99997									
4.5	0.999997									
5.0	0.9999997									

סטטיסטיקה תיאורית

סימנים בסיסיים לשימוש :

אורך האינטרוול	l	אמצע הקבוצה	$or \quad x'$ MP
צפיפות האינטרוול	d	גודל המדגם	N
ממוצע	$\bar{x}$	סכום אברים	Σ
חציון	M_e	שכיחות מוחלטת	f
השכיח	M_0	שכיחות מצטברת	F
סטיית תקן	σ	שכיחות יחסית	p
שונות	σ^2	ציון תקן	z_i

n - מספר קבוצות (מחלקות)

L_i - גבול תחתון של קטע i .

R_i - גבול עליון של קטע i .

$$x'_i = (L_{in} + R_i) / 2$$

מציאת טווח:

$$R = x_n - x_1 \quad (1)$$

(קיבוץ תצפיות לפי ערכים בודדים)

$$R = R_n - L_1 \quad (2)$$

(קיבוץ נתונים לפי קטעים).

$$p_i = \frac{f_i}{N}$$

$$d_i = \frac{f_i}{l_i},$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n x'_i \cdot f_i = \sum_{i=1}^n x'_i \cdot P_i$$

$$\sigma^2 = \frac{1}{N} \sum f_i \cdot (x'_i - \bar{x})^2 = \sum P_i \cdot (x'_i - \bar{x})^2 = \frac{1}{N} \sum f_i \cdot x_i'^2 - (\bar{x})^2 = \sum_{i=1}^n (x'_i)^2 \cdot p_i - (\bar{x})^2$$

$$M_k(X) = \frac{1}{N} \sum_{i=1}^n (x'_i)^k f_i = \sum_{i=1}^n (x'_i)^k \cdot P_i$$

$$\bar{X} = M_1(X)$$

$$\sigma^2 = M_2(X) - (M_1(X))^2$$

מציאת שכיח (קיבוץ תצפיות לפי קטעים)

איחוד של k קבוצות בגודל

$$Mo = L_i + \frac{d_i - d_{i-1}}{(d_i - d_{i-1}) + (d_i - d_{i+1})} \cdot l_i$$

i - קבוצת השכיח $(f_i \geq f_j, 1 \leq j \leq n)$

מציאת חציון (קיבוץ תצפיות לפי קטעים)

$$Me = L_i + \frac{l_i}{f_i} \times (\frac{N}{2} - F_{i-1})$$

i - קבוצת החציון $(F_{i-1} < \frac{N}{2} \text{ ו- } F_i > \frac{N}{2})$.

אם ניתן למצוא את i שעבורו $F_i = \frac{N}{2}$,

$$M_e = R_i$$

מציאת חציון (קיבוץ תצפיות לפי ערכים בודדים)

אם ניתן למצוא את i שעבורו $F_i > \frac{N}{2}$ ו- $F_{i-1} < \frac{N}{2}$.

$$M_e = x_i$$

אם ניתן למצוא את i שעבורו $F_i = \frac{N}{2}$,

$$M_e = (x_i + x_{i+1}) / 2$$

$$z_i = \frac{x'_i - \bar{x}}{\sigma}$$

אקסיומות ההסתברות

(1) תכונות כלליות של פונקציית הסתברות $P(A)$ כאשר A - מאורע כלשהו

Ω - מרחב מאורעות או אוסף כל התוצאות האפשריות של ניסוי או מאורע וודאי.

$$P(\Phi) = 0, P(\Omega) = 1, 0 \leq P(A) \leq 1$$

(2) $P(A) + P(\bar{A}) = 1$, כאשר $\bar{A}$ - מאורע משלים $(A \cup \bar{A} = \Omega)$ $P(\bar{A}) = 1 - P(A)$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (3)$$

(4) A, B מאורעות זרים אם $A \cap B = \Phi$ ומתקיים $P(A \cup B) = P(A) + P(B)$

נוסחאות שימושיות נוספות:

$$P(A \cap \bar{B}) = P(A) - P(A \cap B) \quad (5)$$

$$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}), P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B}) \quad (6) \quad \text{(דה מורגן)}$$

קומבינטוריקה

תמורות

$$P_n = n!; \quad P_n(n_1, n_2, \dots, n_k) = \frac{n!}{n_1! n_2! \dots n_k!}$$

בחירות

	עם החזרה	בלי החזרה
סדר חשוב (חליפות)	n^k	$P_n^k = A_n^k = \frac{n!}{(n-k)!}$
סדר לא חשוב (צירופים)		$C_n^k = \frac{n!}{k!(n-k)!}$

הסתברות מותנית

$P(A/B)$ - קוראים: ההסתברות המותנה של מאורע A כאשר ידוע שהמאורע B התרחש (בקיצור, ההסתברות של A בתנאי B).

$$P(A/B) = \frac{P(A) \cdot P(B/A)}{P(B)}, \quad P(A/B) = \frac{P(A \cap B)}{P(B)}$$

אם A, B מאורעות זרים אז מתקיים $P(A/B) = 0$

אם A, B מאורעות בלתי תלויים אז מתקיים $P(A/B) = P(A)$, $P(A \cap B) = P(A) \cdot P(B)$

אם A, B מאורעות תלויים אז מתקיים $P(A/B) \neq P(A)$, $P(A \cap B) \neq P(A) \cdot P(B)$

נוסחת ההסתברות השלמה:

$$P(A) = P(A/E_1) \cdot P(E_1) + \dots + P(A/E_n) \cdot P(E_n)$$

כאשר $E_1, E_2, \dots, E_n$ מאורעות זרים שאיחודם הוא כל המרחב Ω : $\sum_{i=1}^n P(E_i) = 1$

נוסחת בייז:

$$P(E_1/A) = \frac{P(A/E_1) \cdot P(E_1)}{P(A)} = \frac{P(A/E_1) \cdot P(E_1)}{P(A/E_1) \cdot P(E_1) + \dots + P(A/E_n) \cdot P(E_n)}$$

$$P(\bar{A}/B) = 1 - P(A/B) \quad \text{נוסחה נוספת:}$$

מומנטים למשתנה בדיד

תוחלת $E(x) = \sum x_i P(x_i)$, שונות $\text{Var}(x) = E(x^2) - (E(x))^2$
תכונות של תוחלת ושונות:

$$E(b \cdot X + a) = b \cdot E(X) + a, \quad \text{Var}(b \cdot X + a) = b^2 \cdot \text{Var}(X), \quad E(X + Y) = E(X) + E(Y)$$

נוסחת התוחלת השלמה

נתונים קבוצה שלמה של מאורעות זרים $\{A_1, \dots, A_n\}$. משתנה מקרי כלשהו אזי:
 $E(X) = E(X|A_1) \cdot P(A_1) + E(X|A_2) \cdot P(A_2) + \dots + E(X|A_n) \cdot P(A_n)$

התפלגות בינומית (ברנולי) $X \sim B(n, p)$

n - מספר ניסויים, p - הסתברות של הצלחה בכל ניסוי בודד, $q = 1 - p$ - נסמן:

ההסתברות שמתוך n ניסויים יהיו בדיוק k הצלחות: $P(X = k) = C_n^k p^k q^{n-k}$

תוחלת ושונות של התפלגות בינומית: $E(X) = np$, $VAR(X) = npq$

התפלגות גיאומטרית (עד הצלחה ראשונה) $X \sim G(p)$

p - הסתברות של הצלחה בכל ניסוי בודד, נסמן: $q = 1 - p$

ההסתברות שההצלחה הראשונה תתרחש בניסוי מס' k : $P(X = k) = q^{k-1} \cdot p$

$$E(X) = \frac{1}{p}, \quad VAR(X) = \frac{1-p}{p^2}, \quad P(X > k) = q^k, \quad P(X \leq k) = 1 - q^k$$

$$S = a + aq + aq^2 + \dots = \frac{a}{1-q} \quad (q < 1) \quad \text{סכום של סדרה הנדסית אינסופית}$$

התפלגות בינומית שלילית (מספר ניסיונות עד הצלחה מספר r) $X \sim NB(r, p)$

p - הסתברות של הצלחה בכל ניסוי בודד, נסמן: $q = 1 - p$

ההסתברות לכך שההצלחה מספר r יבוא בניסוי מס' k : $P(X = k) = C_{k-1}^{r-1} p^r q^{k-r}$

$$E(X) = \frac{r}{p}, \quad VAR(X) = \frac{r(1-p)}{p^2}$$

הערה: התפלגות הגיאומטרית הוא מקרה פרטי של התפלגות בינומית שלילית, עם $r = 1$.

נתונים $X_1, X_2, \dots, X_r$ משתנים מקריים גיאומטריים בלתי תלויים כך ש- $X_i \sim G(p)$ עבור כל $1 \leq i \leq r$, אזי $\sum_{i=1}^r X_i \sim NB(r, p)$.

התפלגות היפרגאומטרית

באוכלוסיה N פריטים, מתוכם D פריטים מיוחדים. מוציאים מהאוכלוסיה n פריטים ללא החזרה. X = מספר פריטים מיוחדים שהוצאו.

ההסתברות שמתוך n פריטים יהיו k פריטים מיוחדים: $P(X = k) = \frac{C_D^k \cdot C_{N-D}^{n-k}}{C_N^n}$

תוחלת ושונות:

$$E(X) = n \cdot \frac{D}{N} \quad VAR(X) = n \cdot \frac{D}{N} \cdot \left(1 - \frac{D}{N}\right) \cdot \left(1 - \frac{n-1}{N-1}\right)$$

התפלגות פואסון

התפלגות פואסון עם פרמטר λ : $X \sim P(\lambda)$

λ הוא המספר הממוצע של האירועים ליחידת זמן, $\lambda > 0$

ההסתברות שביחידת הזמן יתרחשו בדיוק k אירועים היא: $P(X = k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}$

תוחלת ושונות: $VAR(X) = \lambda$, $E(X) = \lambda$

קירוב פואסון לבינומית (עבור n גדול ו p קטן) מציבים $\lambda = np$

משתנה מקרי רציף

$f(x)$ פונקציית צפיפות, $a < x < b$, $F(x)$ פונקציית ההתפלגות.

$$F(b) = 1, \quad F(a) = 0, \quad F'(x) = f(x), \quad F(x) = \int_a^x f(t) dt, \quad \int_a^b f(x) dx = 1$$

$F(x)$ אינה יורדת, כלומר אם: $a \leq c < d \leq b$ אז $F(c) \leq F(d)$.

$$P(X < c) = \int_a^c f(x) dx = F(c), \quad P(c < X < d) = \int_c^d f(x) dx = F(d) - F(c)$$

$$E(X) = \int_a^b x \cdot f(x) dx, \quad \text{Var}(X) = E(X^2) - (E(X))^2, \quad E(X^2) = \int_a^b x^2 \cdot f(x) dx,$$

התפלגות מעריכית $X \sim \exp(\lambda)$

פונקצית הצפיפות של X היא: $f(x) = \lambda e^{-\lambda x}$ כאשר $x \geq 0$.
פונקצית ההתפלגות של X היא: $F(t) = P(X \leq t) = 1 - e^{-\lambda t}$.

$$E(X) = \frac{1}{\lambda}, \quad \text{VAR}(X) = \frac{1}{\lambda^2}$$

התפלגות ארלנג $X \sim \text{Erlang}(k, \lambda)$

פונקצית הצפיפות של X היא: $f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!}$

$$E(X) = \frac{k}{\lambda}, \quad \text{VAR}(X) = \frac{k}{\lambda^2}$$

הערה: נתונים $X_1, X_2, \dots, X_k$ משתנים מקריים מעריכיים בלתי תלויים כך ש- $X_i \sim \exp(\lambda)$ עבור כל $1 \leq i \leq k$ אזי $\sum_{i=1}^k X_i \sim \text{Erlang}(k, \lambda)$.

התפלגות נורמלית $X \sim N(\mu, \sigma^2)$ (גאוס)

X הוא משתנה נורמלי עם פרמטרים μ ו- σ , אם פונקצית הצפיפות שלו ניתנת על ידי סימונים: $E(X) = \mu$, $\text{Var}(X) = \sigma^2$.

אם $X \sim N(\mu, \sigma)$ אזי $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

פונקצית ההתפלגות של X היא: $P(X \leq t) = \Phi\left(\frac{t - \mu}{\sigma}\right)$

כאשר $\Phi(t) = P(Z \leq t)$ היא פונקצית ההתפלגות המצטברת של המשתנה הנורמלי הסטנדרטי והיא נתונה בטבלה. תכונות של פונקצית $\Phi(t)$:

$$z_{1-\alpha} = -z_\alpha, \quad P(t_1 < Z < t_2) = \Phi(t_2) - \Phi(t_1), \quad P(-a < Z < a) = 2\Phi(a) - 1$$

$$\Phi(-t) = 1 - \Phi(t), \quad P(Z > t) = 1 - P(Z \leq t)$$

סכום והפרש התפלגויות נורמליות:

נתונים $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$ משתנים מקריים בלתי תלויים, אזי מתקיים:

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

משפט הגבול המרכזי:

היו $X_1, \dots, X_n$ - משתנים מקריים בלתי תלויים בעלי התפלגות זהה שתוחלתה μ ושונותה σ^2 אזי:

$$S \sim N(n \cdot \mu, (\sigma \cdot \sqrt{n})^2) \quad ; \quad S = \sum X_i$$

$$\bar{X} \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) \quad ; \quad \bar{X} = \frac{\sum X_i}{n}$$

קירוב נורמלי למשתנה בינומי:

יהי X משתנה בינומי בעל פרמטרים n ו- p

כלומר $X \sim B(n, p)$. אזי לכל t ממשי עבור n גדול (לפחות: $n > 30$) מתקיים:

$$q = p - 1 ; X \sim N(n \cdot p, \sqrt{n \cdot p \cdot q^2})$$

$$P(k_1 \leq X \leq k_2) = \Phi\left(\frac{k_2 + 1/2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - 1/2 - np}{\sqrt{np(1-p)}}\right)$$

רווח בר- סמך לממוצע כאשר סטיית התקן ידועה:

μ - ממוצע האוכלוסייה, σ^2 - שונות האוכלוסייה, σ - סטיית התקן של האוכלוסייה.
 $\bar{x}$ - ממוצע המדגם, α - רמת המובהקות, $1-\alpha$ רמת הביטחון, n - גודל המדגם, d - סטיית המדגם.

$$\text{רווח בר-סמך ל- } \mu : \bar{x} \pm \frac{\left(z_{1-\frac{\alpha}{2}}\right) \cdot \sigma}{\sqrt{n}}$$

גודל המדגם הדרוש כדי שברמת בטחון $1-\alpha$, הסטייה בין μ לבין $\bar{X}$ (חצי אורך הרווח) לא תעלה על d הוא:

$$d = \frac{\left(z_{1-\frac{\alpha}{2}}\right) \cdot \sigma}{\sqrt{n}} ; n \geq \left\{ \frac{\left(z_{1-\frac{\alpha}{2}}\right) \sigma}{d} \right\}^2$$

רווח בר-סמך ברמת סמך עבור פרופורציה p של "הצלחות" באוכלוסייה:

$$\hat{p} \pm \left(z_{1-\frac{\alpha}{2}}\right) \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

כאשר $\hat{p} = \frac{X}{n}$ היא פרופורציית ה"הצלחות" במדגם בגודל n .

אם ידוע בוודאות ש- $p < \hat{p} < 0.5$ או $p > \hat{p} > 0.5$ אזי

גודל המדגם המבטיח שברמת סמך $1-\alpha$ הסטייה בין p לבין $\hat{p}$ לא תעלה על $d = z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$ והוא:

$$n \geq \frac{z_{1-\alpha/2}^2 \hat{p}\hat{q}}{d^2}$$

אם לא שוללים את האופציה ש- p קרוב ל- 0.5 אזי מתקיים: $n \geq \left(\frac{z_{1-\alpha/2}}{2d}\right)^2$

דף נוסחאות מורחב עם אישור בלבד

סטטיסטיקה תיאורית

סימנים בסיסיים לשימוש :

אורך האינטרוול	l	אמצע הקבוצה	or x' MP
צפיפות האינטרוול	d	גודל המדגם	N
ממוצע	$\bar{x}$	סכום אברים	Σ
חציון	M_e	שכיחות מוחלטת	f
השכיח	M_0	שכיחות מצטברת	F
סטיית תקן	σ	שכיחות יחסית	p
שונות	σ^2	ציון תקן	z_i

n - מספר קבוצות (מחלקות)

L_i - גבול תחתון של קטע i .

R_i - גבול עליון של קטע i .

$$x'_i = (L_{in} + R_i) / 2$$

מציאת טווח:

$$R = x_n - x_1 \quad (1)$$

(קיבוץ תצפיות לפי ערכים בודדים)

$$R = R_n - L_1 \quad (2)$$

(קיבוץ נתונים לפי קטעים.)

$$p_i = \frac{f_i}{N}$$

$$d_i = \frac{f_i}{l_i},$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n x'_i \cdot f_i = \sum_{i=1}^n x'_i \cdot P_i$$

$$\sigma^2 = \frac{1}{N} \sum f_i \cdot (x'_i - \bar{x})^2 = \sum P_i \cdot (x'_i - \bar{x})^2 = \frac{1}{N} \sum f_i \cdot x_i'^2 - (\bar{x})^2 = \sum_{i=1}^n (x'_i)^2 \cdot p_i - (\bar{x})^2$$

מציאת שכיח (קיבוץ תצפיות לפי קטעים)

$$Mo = L_i + \frac{d_i - d_{i-1}}{(d_i - d_{i-1}) + (d_i - d_{i+1})} \cdot l_i$$

i - קבוצת השכיח $(f_i \geq f_j, 1 \leq j \leq n)$

מציאת חציון (קיבוץ תצפיות לפי קטעים)

$$Me = L_i + \frac{l_i}{f_i} \times (\frac{N}{2} - F_{i-1})$$

i - קבוצת החציון $(F_{i-1} < \frac{N}{2} \text{ ו- } F_i > \frac{N}{2})$.

אם ניתן למצוא את i שעבורו $F_i = \frac{N}{2}$,

$$M_e = R_i$$

מציאת חציון (קיבוץ תצפיות לפי ערכים בודדים)

אם ניתן למצוא את i שעבורו $F_i > \frac{N}{2}$ ו- $F_{i-1} < \frac{N}{2}$.

$$M_e = x_i$$

אם ניתן למצוא את i שעבורו $F_i = \frac{N}{2}$,

$$M_e = (x_i + x_{i+1}) / 2$$

$$z_i = \frac{x'_i - \bar{x}}{\sigma}$$

אקסיומות ההסתברות

(1) תכונות כלליות של פונקציית הסתברות $P(A)$ כאשר A - מאורע כלשהו

Ω - מרחב מאורעות או אוסף כל התוצאות האפשריות של ניסוי או מאורע וודאי.

$$P(\Phi) = 0, P(\Omega) = 1, 0 \leq P(A) \leq 1$$

(2) $P(A) + P(\bar{A}) = 1$, $P(\bar{A}) = 1 - P(A)$ כאשר $\bar{A}$ - מאורע משלים $(A \cup \bar{A} = \Omega)$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (3)$$

(4) A, B מאורעות זרים אם $A \cap B = \Phi$ ומתקיים $P(A \cup B) = P(A) + P(B)$

נוסחאות שימושיות נוספות:

$$P(A \cap \bar{B}) = P(A) - P(A \cap B) \quad (5)$$

$$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}), P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B}) \quad (6) \quad (\text{דה מורגן})$$

שיטת ריבוע הקסם:

מציירים ריבוע. בקצה השמאלי כותבים $A, \bar{A}$ ובקצה העליון כותבים $B, \bar{B}$, כך נקבל שתי שורות ושתי עמודות.

בקצה הימני של השורה של A כותבים $P(A)$,

בקצה הימני של השורה של $\bar{A}$ כותבים $P(\bar{A})$,

בקצה התחתון של העמודה של B כותבים $P(B)$,

בקצה התחתון של העמודה של $\bar{B}$ כותבים $P(\bar{B})$.

בפינה הימנית התחתונה כותבים 1, בארבע המשבצות באמצע כותבים את הערכים :

$$P(A \cap B) \text{ בשורה של } A \text{ ובעמודה של } B$$

$$P(\bar{A} \cap B) \text{ בשורה של } A \text{ ובעמודה של } \bar{B}$$

$$P(A \cap \bar{B}) \text{ בשורה של } \bar{A} \text{ ובעמודה של } \bar{B}$$

$$P(\bar{A} \cap \bar{B}) \text{ בשורה של } \bar{A} \text{ ובעמודה של } \bar{B}$$

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

$$P_n = n!; \quad P_n(n_1, n_2, \dots, n_k) = \frac{n!}{n_1! n_2! \dots n_k!}$$

בחירות

	עם החזרה	בלי החזרה
סדר חשוב (חליפות)	n^k	$P_n^k = A_n^k = \frac{n!}{(n-k)!}$
סדר לא חשוב (צירופים)		$C_n^k = \frac{n!}{k!(n-k)!}$

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

$P(A/B)$ - קוראים: ההסתברות המותנה של מאורע A כאשר ידוע שהמאורע B התרחש (בקיזור, ההסתברות של A בתנאי B).

$$P(A/B) = \frac{P(A) \cdot P(B/A)}{P(B)}, \quad P(A/B) = \frac{P(A \cap B)}{P(B)}$$

אם A, B מאורעות זרים אז מתקיים $P(A/B) = 0$

אם A, B מאורעות בלתי תלויים אז מתקיים $P(A/B) = P(A)$, $P(A \cap B) = P(A) \cdot P(B)$

אם A, B מאורעות תלויים אז מתקיים $P(A \cap B) \neq P(A) \cdot P(B)$, $P(A/B) \neq P(A)$

נוסחת ההסתברות השלמה:

$$P(A) = P(A/E_1) \cdot P(E_1) + \dots + P(A/E_n) \cdot P(E_n)$$

כאשר $E_1, E_2, \dots, E_n$ מאורעות זרים שאיחודם הוא כל המרחב Ω : $\sum_{i=1}^n P(E_i) = 1$

נוסחת בייז:

$$P(E_1/A) = \frac{P(A/E_1) \cdot P(E_1)}{P(A)} = \frac{P(A/E_1) \cdot P(E_1)}{P(A/E_1) \cdot P(E_1) + \dots + P(A/E_n) \cdot P(E_n)}$$

$$P(\bar{A}/B) = 1 - P(A/B) \quad \text{נוסחה נוספת:}$$

שיטת העצים:

בונים עץ כך שבהתחלה משרטטים n ענפים מסומנים ב- $E_1, E_2, \dots, E_n$ שיוצאים מנקודה מסוימת, וכותבים על כל ענף את ההסתברויות המתאימות $P(E_1), P(E_2), \dots, P(E_n)$. אחר כך, מורידים מקצה של כל אחד מהענפים ענף נוסף המסומן ב- $A|E_1, A|E_2, \dots, A|E_n$ בהתאמה, וכותבים על כל אחד מהענפים החדשים את ההסתברויות $P(A|E_1), P(A|E_2), \dots, P(A|E_n)$. כעת, כדי לחשב $P(A)$, מכפילים כל אחד מהערכים $P(A|E_1), P(A|E_2), \dots, P(A|E_n)$ בערכים $P(E_1), P(E_2), \dots, P(E_n)$ שמעליו, ומחברים את הערכים מהענפים השונים שהתקבלו.

מומנטים למשתנה בדיד

$$E(x) = \sum x_i P(x_i), \quad \text{שונות} \quad \text{תוחלת} \quad \text{שונות} \quad \text{תוחלת של תוחלת ושונות:}$$

$$E(b \cdot X + a) = b \cdot E(X) + a, \quad \text{Var}(b \cdot X + a) = b^2 \cdot \text{Var}(X), \quad E(X + Y) = E(X) + E(Y)$$

התפלגות בינומית (ברנולי) $X \sim B(n, p)$

n - מספר ניסויים, p - הסתברות של הצלחה בכל ניסוי בודד, $q = 1 - p$: נסמן

הסתברות שמתוך n ניסויים יהיו בדיוק k הצלחות: $P(X = k) = C_n^k p^k q^{n-k}$

תוחלת ושונות של התפלגות בינומית: $E(X) = np$, $\text{VAR}(X) = npq$

"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 מהאוכלוסיה n פריטים **ללא החזרה**. X = מספר פריטים מיוחדים שהוצאו.

הסתברות שמתוך n פריטים יהיו k פריטים מיוחדים: $P(X = k) = \frac{C_D^k \cdot C_{N-D}^{n-k}}{C_N^n}$

תוחלת ושונות:

$$E(X) = n \cdot \frac{D}{N} \quad \text{VAR}(X) = n \cdot \frac{D}{N} \cdot \left(1 - \frac{D}{N}\right) \cdot \left(1 - \frac{n-1}{N-1}\right)$$

התפלגות גיאומטרית (עד הצלחה ראשונה) $X \sim G(p)$

p - הסתברות של הצלחה בכל ניסוי בודד, נסמן: $q = 1 - p$

ההסתברות שההצלחה הראשונה תתרחש בניסוי מס' k : $P(X = k) = q^{k-1} \cdot p$

$$E(X) = \frac{1}{p}, \quad \text{VAR}(X) = \frac{1-p}{p^2}, \quad P(X > k) = q^k, \quad P(X \leq k) = 1 - q^k$$

$$S = a + aq + aq^2 + \dots = \frac{a}{1-q} \quad (q < 1) \quad \text{סכום של סדרה הנדסית אינסופית}$$

התפלגות פואסון

התפלגות פואסון עם פרמטר λ : $X \sim P(\lambda)$

λ הוא המספר הממוצע של האירועים ליחידת זמן, $\lambda > 0$

$$P(X = k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!} \quad \text{ההסתברות שביחידת הזמן יתרחשו בדיוק } k \text{ אירועים היא}$$

$$\text{VAR}(X) = \lambda, \quad E(X) = \lambda \quad \text{תוחלת ושונות}$$

קירוב פואסון לבינומית (עבור n גדול ו p קטן) מציבים $\lambda = np$

בהתפלגות פואסון חשוב להשתמש במוצא ליחידת הזמן שרוצים לחשב, ואם בנתון המקורי של הממוצע הוא ליחידת 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 λ שיהיה מספר הפעמים שהאירוע יתרחש מספר הפעמים בממוצע בשעה).

משתנה מקרי רציף

$f(x)$ פונקציית צפיפות, $a < x < b$, $F(x)$ פונקציית ההתפלגות.

$$F(b) = 1, \quad F(a) = 0, \quad F'(x) = f(x), \quad F(x) = \int_a^x f(t) dt, \quad \int_a^b f(x) dx = 1$$

$F(x)$ אינה יורדת, כלומר אם: $a \leq c < d \leq b$ אז $F(c) \leq F(d)$.

$$P(X < c) = \int_a^c f(x) dx = F(c), \quad P(c < X < d) = \int_c^d f(x) dx = F(d) - F(c)$$

$$E(X) = \int_a^b x \cdot f(x) dx, \quad \text{Var}(X) = E(X^2) - (E(X))^2, \quad E(X^2) = \int_a^b x^2 \cdot f(x) dx,$$

התפלגות נורמלית (גאוס) $X \sim N(\mu, \sigma^2)$

X הוא משתנה נורמלי עם פרמטרים μ ו- σ , אם פונקציית הצפיפות שלו ניתנת על ידי

$$\text{סימונים: } E(X) = \mu, \quad \text{Var}(X) = \sigma^2$$

$$\text{אם } X \sim N(\mu, \sigma) \text{ אזי } Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

$$\text{פונקציית התפלגות של } X \text{ היא: } P(X \leq t) = \Phi\left(\frac{t - \mu}{\sigma}\right)$$

כאשר $\Phi(t) = P(Z \leq t)$ היא פונקציית ההתפלגות המצטברת של המשתנה הנורמלי הסטנדרטי

והיא נתונה בטבלה. תכונות של פונקציית $\Phi(t)$:

$$P(-a < Z < a) = 2\Phi(a) - 1, \quad P(t_1 < Z < t_2) = \Phi(t_2) - \Phi(t_1), \quad z_{1-\alpha} = -z_\alpha$$

$$\Phi(-t) = 1 - \Phi(t), \quad P(Z > t) = 1 - P(Z \leq t)$$

משפט הגבול המרכזי

יהיו $X_1, \dots, X_n$ - משתנים מקריים בלתי תלויים בעלי התפלגות זהה שתוחלתה μ ושונותה σ^2 אזי

$$P\left(\sum_{i=1}^n X_i \leq t\right) = \Phi\left(\frac{t - n\mu}{\sigma\sqrt{n}}\right) \quad \text{או} \quad \sum_{i=1}^n X_i \sim N(n\cdot\mu, n\cdot\sigma^2)$$

$$P\left(\bar{X} \leq t\right) = \Phi\left(\frac{t - \mu}{\sigma/\sqrt{n}}\right) \quad \text{או} \quad \bar{X} \sim N(\mu, \sigma^2/n)$$

קירוב נורמלי למשתנה בינומי

יהי X משתנה בינומי בעל פרמטרים n ו- p כלומר $X \sim B(n, p)$, אזי עבור n מספיק גדול ($n > 30$) מתקיים:

$$X \sim N(np, np(1-p)); P(X \leq k) = \Phi\left(\frac{k + 1/2 - np}{\sqrt{np(1-p)}}\right)$$

סכום והפרש התפלגויות נורמליות

יהיו X ו- Y משתנים מקריים נורמליים בלתי תלויים כך ש: $X \sim N(\mu_x, \sigma_x^2); Y \sim N(\mu_y, \sigma_y^2)$

אזי מתקיים: $X \pm Y \sim N(\mu_x \pm \mu_y, \sigma_x^2 + \sigma_y^2)$

$$\bar{X} - \bar{Y} \sim N\left(\mu_x - \mu_y, \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}\right); P(\bar{X} - \bar{Y} \leq t) = \Phi\left(\frac{t - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}\right)$$

רווח בר-סמך לממוצע כאשר סטיית התקן ידועה:

μ - ממוצע האוכלוסייה, σ^2 - שונות האוכלוסייה, σ - סטיית התקן של האוכלוסייה.

$\bar{x}$ - ממוצע המדגם, α - רמת המובהקות, $1-\alpha$ רמת הביטחון, n - גודל המדגם

$$\bar{x} - \frac{z_{1-\alpha/2} \cdot \sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{z_{1-\alpha/2} \cdot \sigma}{\sqrt{n}}; \quad \Phi(z_{1-\alpha/2}) = 1 - \frac{\alpha}{2} \quad : \mu \text{ - רווח בר-סמך ל-}$$

גודל המדגם הדרוש כדי שברמת בטחון $1-\alpha$, הסטייה בין μ לבין $\bar{X}$ לא תעלה על $d = \frac{z_{1-\alpha/2} \sigma}{\sqrt{n}}$ הוא:

$$n \geq \left(\frac{z_{1-\alpha/2} \sigma}{d}\right)^2$$

רווח בר-סמך עבור פרופורציה

$$\hat{p} - z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad : \text{רווח בר-סמך עבור פרופורציה } p$$

כאשר $\hat{p} = \frac{x}{n}$ היא פרופורציית ה"הצלחות" במדגם בגודל n

אם ידוע בוודאות ש- $p < \hat{p} < 0.5$ או $p > \hat{p} > 0.5$ אזי

גודל המדגם המבטיח שברמת סמך $1-\alpha$ הסטייה בין p לבין $\hat{p}$ לא תעלה על $d = z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

והוא : $n \geq \frac{z_{1-\alpha/2}^2 \cdot \hat{p} \cdot (1-\hat{p})}{d^2}$ כאשר $\hat{p}$ הוא הערך הקרוב ביותר ל-0.5 שיכול להתקבל.

אם לא שוללים את האופציה ש- p קרוב ל-0.5 אזי מתקיים : $n \geq \left(\frac{z_{1-\alpha/2}}{2d} \right)^2$

אינטגרציה וגזירה בחלקים:

$$[u(x)v(x)]' = u'(x)v(x) + v'(x)u(x)$$

$$[u(x)/v(x)]' = [u'(x)v(x) - v'(x)u(x)]/v^2(x)$$

$$\int u'(x)v(x)dx = u(x)v(x) - \int v'(x)u(x)dx$$