


הסתברות

התפלגות בינומית ונורמלית


משתנים מקריים – משתנה ברנולי

- ניסוי ברנולי הוא ניסוי מקרי שבו יש בדיוק שני מאורעות יסודיים: הצלחה וכישלון.
- ההסתברות להצלחה היא בניסוי היא p, וההסתברות לכישלון היא $q = 1 - p$
- נגדיר משתנה מקרי X השווה ל-1 אם הניסוי הסתיים בהצלחה, ושווה ל-0 אם הניסוי הסתיים בכישלון, כלומר:

$$P(X = 1) = p$$

$$P(X = 0) = 1 - p$$

- קל לחשב כי

$$E(X) = p$$

$$V(X) = p \cdot (1 - p)$$

משתנה מקרי בינומי

- עורכים סדרה של n ניסויי ברנולי זהים ובלתי תלויים עם הסתברות להצלחה p. נסמן ב-X את מספר הניסויים שהסתיימו בהצלחה
- הערכים ש-X יכול לקבל הם: $k=0,1,2,\ldots,n$
- המשתנה X נקרא בשם משתנה מקרי בינומי. מסמנים: $X\sim B(n,p)$
- X הוא למעשה סכום של n משתני ברנולי עם הסתברות להצלחה p : $X=X_1+X_2+\ldots+X_n$

$$E(X)=E(X_1)+E(X_2)+\ldots+E(X_n)=p+p+\cdots+p=n\cdot p$$

$$V(X)=V(X_1)+V(X_2)+\ldots+V(X_n)=p\cdot(1-p)+\cdots+p\cdot(1-p)+p\cdot(1-p)=n\cdot p\cdot(1-p)$$

- פונקציית ההסתברות של X היא $P(X=k)=\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}$ כאשר $\binom{n}{k}=\frac{n!}{k!\cdot(n-k)!}$

משתנה בינומי - דוגמה

סטודנט עומד לבחינה בה יש 10 שאלות "אמריקאיות", כאשר לכל שאלה יש ארבע תשובות אפשריות ורק תשובה אחת נכונה. הסטודנט לא למד כל הסמסטר, לא הגיש תרגילים ולא התכונן לבחינה, ולכן הוא מנחש את התשובות באופן מקרי לחלוטין. כדי לעבור את הבחינה צריך לקבל ציון 70 לפחות. מה ההסתברות כי הסטודנט יצליח לעבור את הבחינה?

תשובה: מאחר והניחושים הם מקריים לחלוטין, ניתן להניח כי כל שאלה היא ניסוי ברנולי עם הסתברות 0.25 להצלחה. לכן מספר התשובות הנכונות הוא משתנה מקרי בינומי: $X\sim B(10,0.25)$

כדי להצליח בבחינה על הסטודנט לענות נכונה על 7 שאלות לפחות, ולכן עלינו לחשב את $P(X\geq 7)$:

$$P(X\geq 7)=P(X=7)+P(X=8)+P(X=9)+P(X=10)$$

חישוב ההסתברות להצלחה בבחינה

- חישוב: ניתן לבצע, למשל, בעזרת נייר ועט, מחשבון, אקסל (עם הפונקציות BINOM.DIST או BINOM.DIST.RANGE) תוכנות אחרות, מחשבונים באינטרנט:

$$P(X=7) = \binom{10}{7} \cdot 0.25^{7} \cdot 0.75^{8} = 0.00309$$

$$P(X=8) = \binom{10}{8} \cdot 0.25^{8} \cdot 0.75^{2} = 0.00037$$

$$P(X=9) = \binom{10}{9} \cdot 0.25^{9} \cdot 0.75^{1} = 0.00003$$

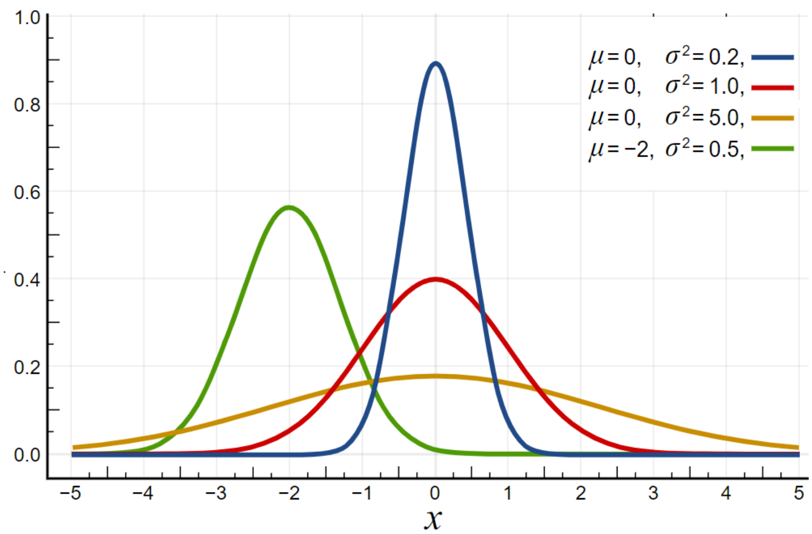
$$P(X=10) = \binom{10}{10} \cdot 0.25^{10} \cdot 0.75^{0} = 0.00000$$

$$P(X \geq 7) = 0.00351)$$

משתנה מקרי נורמלי

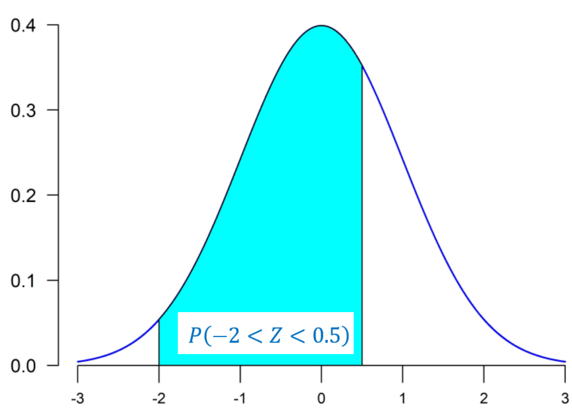
- עד כה ראינו משתנים מקריים בדידים
- המשתנה המקרי הנורמלי (או ההתפלגות הנורמלית) הוא משתנה מקרי רציף
- ההסתברויות שוות לשטח תחת עקומת הצפיפות **(מייד יוסבר בהרחבה)**
- ניתן לחשב את ההסתברויות רק עבור תחומים של ערכים ולא עבור ערכים יחידים
- להתפלגות הנורמלית יש שני פרמטרים: μ ו-σ
- התוחלת של ההתפלגות הנורמלית שווה ל-μ וסטיית התקן שווה ל-σ
- אם μ=0 ו- σ=1 אז ההתפלגות נקראת התפלגות נורמלית סטנדרטית

משתנה מקרי נורמלי

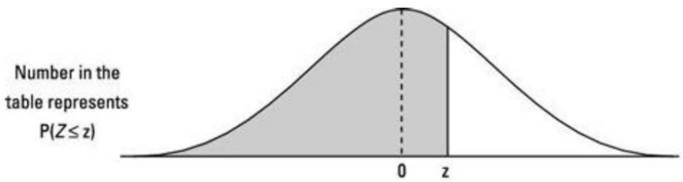


התפלגות נורמלית סטנדרטית

- התפלגות נורמלית סטנדרטית היא התפלגות נורמלית עם תוחלת השווה ל-0 וסטיית תקן השווה ל-1
- נהוג לסמן אותו באות Z
- ההסתברות מבוטאת על ידי שטח מתחת לעקומה
- ניתן לחשב את השטח בעזרת טבלה, אקסל, ויש גם מחשבונים ברשת



דוגמא לחישוב בעזרת טבלה

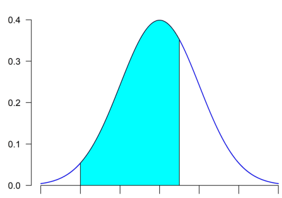


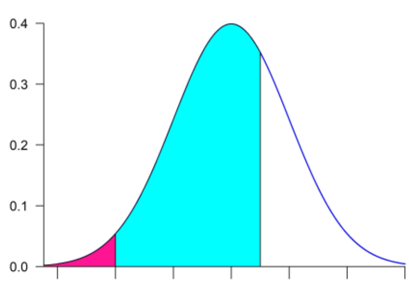
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621

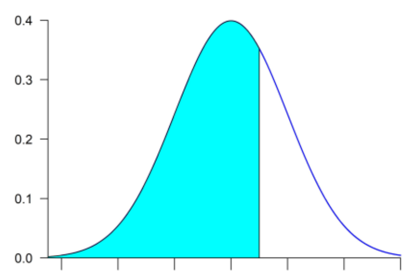
דוגמת חישוב בעזרת טבלה

- דוגמא: רוצים לחשב את $P(-2 < Z < 0.5)$
- אסטרטגיה

1. נחשב את $P(Z < 0.5)$
2. נחסיר את $P(Z < -2)$

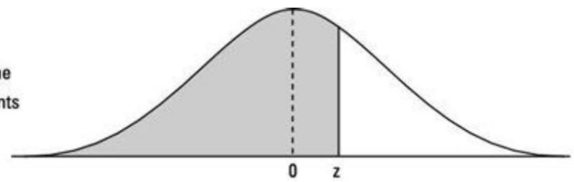






חישוב $P(Z<0.5)$

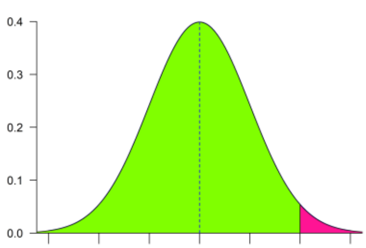
Number in the table represents $P(Z \le z)$

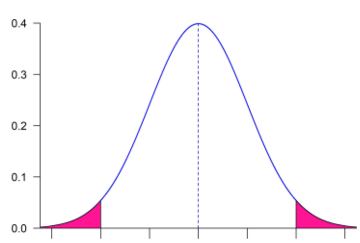


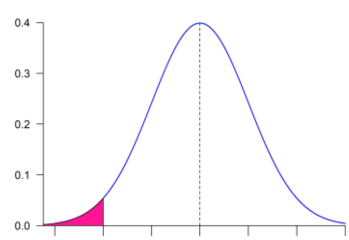
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.6844	.6879
0.5	.6915	.6950	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.7190	.7224
0.6	.7257	.7291	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621

$$P(Z < 0.5) = 0.6915$$

חישוב $P(Z<-2)$







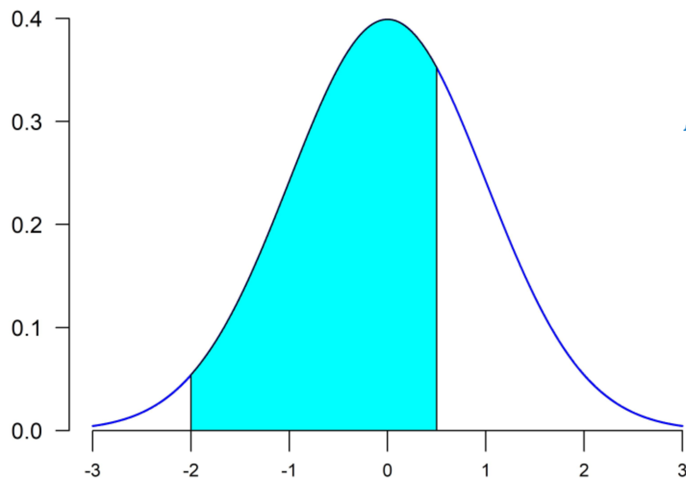
$$P(Z < 2) = 0.9772$$

$$P(Z < -2) = P(Z > 2) \ = 0.0228$$

$$P(Z > 2) = 1 - P(Z < 2)$$

$$= 1 - 0.9772 = 0.0228$$

ובסופו של דבר



$$P(-2 < Z < 0.5) =$$

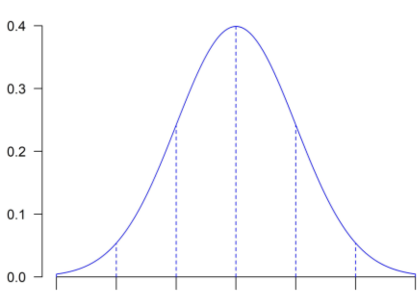
$$P(Z < 0.5) - P(Z < -2) =$$

$$0.6915 - 0.0028 = 0.6687$$

התפלגות נורמלית "רגילה"

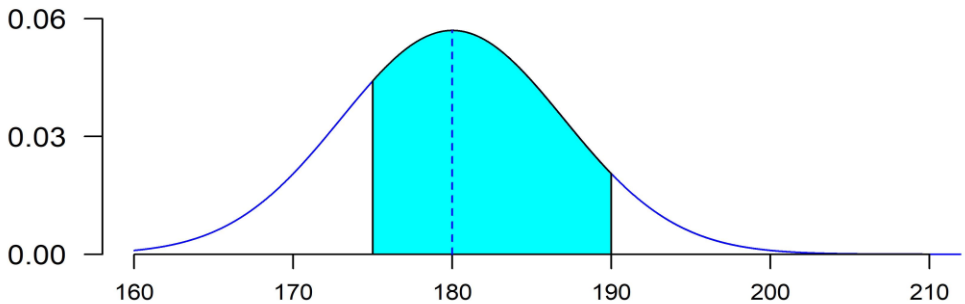
- להתפלגות הנורמלית יש שני פרמטרים: μ ו-σ
- התוחלת של ההתפלגות הנורמלית שווה ל-μ וסטיית התקן שווה ל-σ
- אם המשתנה X מתפלג נורמלית עם תוחלת השווה ל-μ וסטיית תקן השווה ל-σ

אז למשתנה $Z = \dfrac{X-\mu}{\sigma}$ יש התפלגות נורמלית סטנדרטית



דוגמה: התפלגות נורמלית

- במדינה מסויימת גובה הגברים מתפלג נורמלית עם תוחלת 180 ס"מ וסטיית תקן 7 ס"מ. בוחרים גבר אחד באופן מקרי מהאוכלוסייה. מה ההסתברות שגובהו בין 175 ל-190 ס"מ?



פתרון

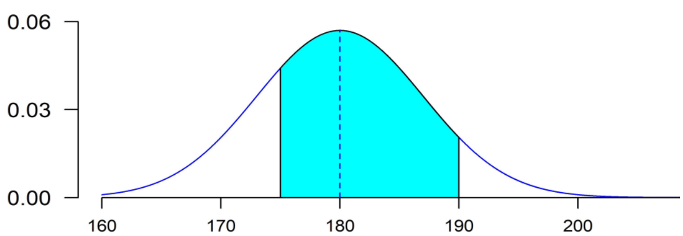
- "ננרמל" את כל הערכים על ידי הפחתת התוחלת – 180, ולאחר מכן בחלוקה בסטיית התקן - 7

$$P(175 < X < 190) \;= P(175 - 180 < X - 180 < 190 - 180)$$

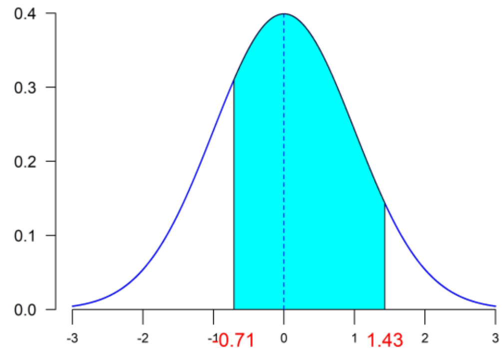
$$= P\left(\frac{175-180}{7} < \frac{X-180}{7} < \frac{190-180}{7}\right) \;= P\left(-0.71 < \frac{X-180}{7} < 1.43\right)$$

$$= P(-0.71 < Z < 1.43)$$

בסופו של דבר



$$P(175 < X < 190)$$



$$= P(-0.71 < Z < 1.43)$$

ממוצע של משתנים מקריים נורמליים

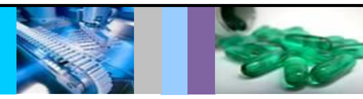
- נתון מדגם מהתפלגות נורמלית עם תוחלת μ וסטיית תקן σ
- המדגם הוא למעשה סדרה של משתנים מקריים מהתפלגות זו, כאשר המשתנים המקריים הם **בלתי תלויים**
- נאמר כי $X_1, X_2, \dots, X_n$ הם "משתנים מקריים בלתי תלויים ושווי התפלגות מהתפלגות נורמלית עם תוחלת μ וסטיית תקן σ"
- נסמן את ממוצע המדגם ב- $\bar{X}_n$:

$$\bar{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

- לפי הכללים של התוחלת והשונות נוכל למצוא כי

$$E(\bar{X}_n) = \mu \qquad V(\bar{X}_n) = \frac{\sigma^2}{n} \qquad SD(\bar{X}_n) = \frac{\sigma}{\sqrt{n}}$$

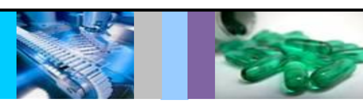
ממוצע של משתנים מקריים נורמליים



ניתן לטעון טענה חזקה יותר לגבי הממוצע של המדגם:

- אם $X_1, X_2, \dots, X_n$ הם משתנים מקריים בלתי תלויים ושווי התפלגות **מהתפלגות נורמלית** עם תוחלת μ וסטיית תקן σ, אז ל- יש **התפלגות נורמלית** עם תוחלת μ וסטיית תקן $\frac{\sigma}{\sqrt{n}}$.
- תכונה זו מאפשרת לנו לחשב הסתברויות ובהמשך להסיק מסקנות סטטיסטיות גם בעזרת **מדגמים קטנים**, אם ההתפלגות היא נורמלית.

דוגמה: בקרת איכות



חברת תרופות מייצרת תרופה מאוד יקרה, ואצוות הייצור הן קטנות. כדי לוודא שכמות החומר הפעיל בתרופה מתאימה לתקן, יש צורך בבדיקה ביולוגית, אבל הבדיקה הורסת את התרופה, ולכן יש צורך במדגם קטן. ידוע כי כמות החומר הפעיל בתרופה שיוצרה בתהליך ייצור תקין מפלגת התפלגות נורמלית עם תוחלת 40 מ"ג וסטיית תקן של 3 מ"ג. תרופה נחשבת תקינה אם כמות החומר הפעיל היא בין 38 ל-42 מ"ג.

מה ההסתברות כי במדגם של 10 תרופות ממוצע כמות החומר הפעיל תהיה בין 38 ל-42 מ"ג?

תשובה: לממוצע המדגם יש **התפלגות נורמלית** עם תוחלת השווה ל-40 וסטיית תקן השווה $\frac{3}{\sqrt{10}} \approx 0.949$

בעזרת טבלה, אקסל או מחשבון התפלגות נורמלית אפשר לחשב כי ההסתברות היא בערך 0.948