הפרמטר	התנאים	דחייה אם: $H_0: \theta = \theta_0$, $H_1: \theta > \theta_0$	דחייה אם: $H_0: \theta = \theta_0$, $H_1: \theta < \theta_0$	דחייה אם: $H_0: \theta = \theta_0$, $H_1: \theta \neq \theta_0$	התקן	התפלגות האומד:	רווח סמך ו-SE	שונות
ממוצע -יחיד μ	1. $x \sim N$ או $n \ge 30$ 2. σ^2 ידועה	$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ **או** $\bar{X} < \mu_0 - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$	$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$	$\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$	$\bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$, $S.E = \frac{\sigma}{\sqrt{n}}$	$n \ge \left(\frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1}\right)^2$ $n \ge \left(\frac{2 \cdot z_{1-\frac{\alpha}{2}} \cdot \sigma}{d_0}\right)^2$
	1. $n \ge 30$ 2. σ^2 אינה ידועה	$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$ **או** $\bar{X} < \mu_0 - Z_{1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$	$Z = \frac{\bar{x} - \mu}{S / \sqrt{n}}$	$\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$	$\bar{x} \pm z_{1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$ $S.E \approx \frac{S}{\sqrt{n}}$	
	1. $x \sim N$ ו $n < 30$ 2. σ^2 אינה ידועה	$\bar{X} > \mu_0 + t^{n-1}{}_{1-\alpha} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} < \mu_0 - t^{n-1}{}_{1-\alpha} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + t^{n-1}{}_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$ **או** $\bar{X} < \mu_0 - t^{n-1}{}_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$	$T = \frac{\bar{x} - \mu}{S / \sqrt{n}}$	$\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$	$\bar{x} \pm t^{(n-1)}_{1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$ $S.E \approx \frac{S}{\sqrt{n}}$	דרגות חופש=n-1
פרופורציה במדגם יחיד p	התנאי: $np \ge 5$ וגם $n(1-p) \ge 5$	$\hat{p} > p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$\hat{p} < p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$\hat{p} < p_0 - Z_{1-\alpha/2} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ **או** $\hat{p} > p_0 + Z_{1-\alpha/2} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$	$\hat{p} \sim N(p, \frac{pq}{n})$	$\hat{p} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$ $S.E = \sqrt{\frac{pq}{n}} \approx \sqrt{\frac{\hat{p}\hat{q}}{n}} \le \sqrt{\frac{1}{4n}}$	$n \ge \left(\frac{Z_{1-\alpha}\sqrt{p_0 q_0} + Z_{1-\beta}\sqrt{p_1 q_1}}{p_0 - p_1}\right)^2$ $n \ge \left(\frac{Z_{1-\frac{\alpha}{2}}}{d_0}\right)^2$
הפרש ממוצעים במדגמים בלתי תלויים- $\mu_1 - \mu_2$ כאשר $H_0:$ $\mu_1 - \mu_2 = C$	1. σ_2^2, σ_1^2 ידועות. 2. $X_1, X_2 \sim N$ או $n_1, n_2 > 30$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ **או** $\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$z = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$\bar{x}_1 - \bar{x}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$	$(\bar{x}_1 - \bar{x}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}$ $S.E = \sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}$	
	1. σ_2^2, σ_1^2 אינן ידועות 2. $n_1 \ge 30$ $n_2 \ge 30$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ **או** $\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$z = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	$\bar{x}_1 - \bar{x}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$	$(\bar{x}_1 - \bar{x}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{s^2{}_1}{n_1} + \frac{s^2{}_2}{n_2}}$	
	1. התפלגות נורמלית 2. $n_1 < 30$ $n_2 < 30$ 3. סטיות התקן אינן ידועות אך $\sigma_1^2 = \sigma_2^2$	$\bar{x}_1 - \bar{x}_2 > c + t_{1-\alpha}^{n_1+n_2-2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$\bar{x}_1 - \bar{x}_2 < c - t_{1-\alpha}^{n_1+n_2-2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + t_{1-\alpha/2}^{n_1+n_2-2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ **או** $\bar{x}_1 - \bar{x}_2 < c - t_{1-\alpha/2}^{n_1+n_2-2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$T = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$\bar{x}_1 - \bar{x}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$	$(\bar{x}_1 - \bar{x}_2) \pm t^{n_1+n_2-2}{}_{1-\frac{\alpha}{2}} \cdot S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ $S.E \approx S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$ דרגות חופש= $n_1 + n_2 - 2$
הפרש פרופורציות במדגמים בלתי תלויים: $H_0: p_1 - p_2 = 0$		$\hat{p}_1 - \hat{p}_2 > 0 + Z_{1-\alpha} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$	$\hat{p}_1 - \hat{p}_2 < 0 - Z_{1-\alpha} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$	$\hat{p}_1 - \hat{p}_2 > 0 + Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$ **או** $\hat{p}_1 - \hat{p}_2 < 0 - Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$	$Z_{\hat{p}_1 - \hat{p}_2} = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$ $Z_{H_0} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$	$\hat{p}_1 - \hat{p}_2 \sim N(p_1 - p_2, \frac{p_1 \cdot q_1}{n_1} + \frac{p_2 \cdot q_2}{n_2})$	$(\hat{p}_1 - \hat{p}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$ $S.E \approx \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$	$\hat{p} = \frac{y_1 + y_2}{n + n} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$
תוחלת ההפרש במדגמים מזווגים. $H_0: \mu_D = C$	1. המדגם מזווג 2. D מתפלג נורמלית.	$\bar{D} > C + t^{n-1}{}_{1-\alpha} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} < C - t^{n-1}{}_{1-\alpha} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} > C + t^{n-1}{}_{1-\frac{\alpha}{2}} \cdot \frac{S_D}{\sqrt{n}}$ **או** $\bar{D} < C - t^{n-1}{}_{1-\frac{\alpha}{2}} \cdot \frac{S_D}{\sqrt{n}}$	$T_{\bar{D}} = \frac{\bar{D} - \mu_D}{S_D / \sqrt{n}}$	$\bar{D} \sim N(\mu_D, \frac{\sigma^2{}_d}{n})$	$\bar{D} \pm t^{n-1}{}_{1-\frac{\alpha}{2}} \frac{S_D}{\sqrt{n}}$ $S.E \approx \frac{S_D}{\sqrt{n}}$	דרגות חופש=n-1

התפלגות ממוצע המדגם $\bar{X}$

א. תכונות: $E(\bar{x})=\mu$ - תוחלת ממוצע המדגם שווה לממוצע האוכלוסייה- כלומר ממוצע של כל ממוצעי המדגם האפשריים נותן את ממוצע האוכלוסייה האמיתי – הפרמטר.

$v(\bar{x})=\frac{v(x)}{n}=\frac{\sigma^2}{n}$ - השונות של ממוצעי המדגם שווה לשונות האוכלוסייה לחלק ל n – גודל המדגם. סטיית התקן: $\sigma(\bar{x})=\frac{\sigma}{\sqrt{n}}$

ב. משפט: אם האוכלוסיה מתפלגת נורמלית $x\sim N(\mu,\sigma^2)$ ונדגום n תצפיות אז ממוצע המדגם גם מתפלג נורמלית $\bar{x}\sim N(\mu,\frac{\sigma^2}{n})$ ואז: $Z=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$

ג.משפט הגבול המרכזי אם x מתפלג כלשהו וידוע $E(x)=\mu$, $V(x)=\sigma^2$ אזי עבור מדגם מספיק גדול (לפחות 30) $\bar{x}\sim\succ N(\mu,\frac{\sigma^2}{n})$

התפלגות סכום התצפיות במדגם

א. תכונות: $E(\sum x_i)=n\mu$ $V(\sum x_i)=n\sigma^2$

ב. משפט: אם $x\sim N(\mu,\sigma^2)$ אזי $\sum x_i\sim N(n\mu,n\sigma^2)$ ו- $Z=\frac{\sum x_i-n\mu}{\sqrt{n\sigma^2}}$

ג. משפט הגבול המרכזי : אם x מתפלג כלשהו וידוע $E(x)=\mu$, $V(x)=\sigma^2$ אזי עבור מדגם מספיק גדול (לפחות 30) $\sum x_i\sim\succ N(n\mu,n\sigma^2)$

התפלגות מספר ההצלחות במדגם-התפלגות בינומית Y:

הסיכוי להצלחה p (גודל זה הוא פרמטר המתאר את האוכלוסייה) הסיכוי לכישלון q=1-p

y~B(n,p)

$E(Y)=np$ $\quad p(y=k)=\binom{n}{k}p^kq^{n-k}$

$V(Y)=npq$ $\quad \binom{n}{k}=\frac{n!}{k!(n-k)!}$

$\sigma(Y)=\sqrt{npq}$

רמת בטחון = רמת סמך: 1-α
$\bar{X}$ - נופל באמצע רווח סמך -רווח סמך A<μ<B
אורך רווח סמך : d=B-A
רמת מובהקות : α

קירוב נורמלי להתפלגות הבינומי:

אם y~B(n,p) והתנאי: $np\geq5$ וגם $n(1-p)\geq5$ מתקיים אז: $y\sim\succ N(np,npq), Z_y=\frac{y-np}{\sqrt{npq}}$ **לא לשכוח תיקון רציפות**

התפלגות פרופורציית ההצלחות במדגם: $\hat{p}=\frac{y}{n}$

$E(\hat{P})=p, V(\hat{P})=\frac{pq}{n}$ $\quad \sigma(\hat{P})=\sqrt{\frac{p\cdot(1-p)}{n}}=SE\leq\sqrt{\frac{1}{4n}}$

אם באוכלוסייה פרופורציית ההצלחות היא P ו $np\geq5$ וגם $n(1-p)\geq5$ נאמר שהתפלגות פרופורציית ההצלחות במדגם $\hat{p}$ היא נורמלית:

$\hat{p}\sim N(p,\frac{pq}{n})$ ו- $Z=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$

התפלגות – גיאומטרית : x – מס' הנסיונות עם ההצלחה ה 1
$X\sim G(p)$ $\quad p(x=k)=(1-p)^{k-1}\cdot p$
$p(x>k)=(1-p)^k$

חוק המספרים הגדולים(כל מה שנאמר על פרופורציה נכון גם לגבי ממוצע):

חוק המספרים הגדולים מתייחס להשפעת גודל המדגם על הסיכוי של פרופורציית המדגם להיות קרובה או רחוקה מהפרופורציה האמיתית:

- ככל שהמדגם גדול יותר כך הסיכוי שפרופורציית המדגם ($\hat{p}$) תהיה בקרבת הפרופורציה באוכלוסייה (P) גבוהה יותר.
- ואילו הסיכוי לקבל ערך חריג הרחוק מהפרופורצייה של האוכלוסייה קטן יותר .
- אולם הסיכוי שפרופורציית המדגם תהיה בדיוק הפרופורצייה האמיתית הולכת וקטנה ככל שהמדגם גדול יותר.

P_v=P(לקבל את תוצאות המדגם וקיצוני | H_0 נכונה)
אם $P_v \geq \alpha$ דוחים את H_0.
הערה: במבחן דו צדדי יש להכפיל את הסיכוי פי 2 כדי לקבל P_v.

- מובהקות התוצאה

$S^2=\frac{\sum(x_i-\bar{x})^2}{n-1}=\frac{\sum x_i^2-n\bar{x}^2}{n-1}$ - שונות המדגם (אומד חסר הטייה ל- σ^2)

הגדרות:

- **פרמטר** הוא גודל קבוע המאפיין את האוכלוסייה (כמו μ) ואילו **סטטיסטי** הוא משתנה המחושב על תוצאות המדגם (כמו $\bar{x}$)
- אם התוחלת של $\hat{\theta}$ תהיה שווה ל θ: $\hat{\theta}$ יהיה **אומד חסר הטייה** ל θ כלומר $E(\hat{\theta})=\theta$
- **טעות התקן** של אומד היא סטיית התקן שלו , כלומר: $\sigma(\hat{\theta})=S.E$
- **שגיאת אמידה** הינו ההפרש בערך מוחלט בין האומד לפרמטר כלומר $|\hat{\theta}-\theta|$
- **רמת סמך = 1-α** הסיכוי שהפרמטר נופל ברווח הסמך כלומר: $P(A\leq\theta\leq B)=1-\alpha$
- **טעות מסוג ראשון** - להכריע לדחות את H_0 למרות שבמציאות H_0 נכונה.
- **טעות מסוג שני** - להכריע לקבל את H_0 למרות שבמציאות H_1 נכונה.

α=**רמת מובהקות**=P(**טעות מסוג ראשון**)=P(H_0 לדחות את \| H_0 נכונה)

(1-α) = רמת בטחון=P(H_0 לקבל את \| H_0 נכונה)

β=P(**טעות מסוג שני**)=P(H_0 לקבל את \| H_1 נכונה)

π=**(1-β)=עוצמה**=P(H_0 לדחות את \| H_1 נכונה)

מבחן טיב התאמה: מבחן שמטרתו לבדוק האם האוכלוסייה מתפלגת לפי התאמה מסוימת.

השערות: H_0: ישנה התאמה מסוימת

H_1: אחרת

כלל הכרעה: $c-1$ דרגות החופש: כאשר c זה מספר הקטגוריות. הערך הקריטי הוא: $k=\chi^{2(c-1)}_{1-\alpha}$ נידחה את השערת האפס אם הערך הסטטיסטי יהיה גדול מהערך הקריטי.

סטטיסטי המבחן: $\boxed{\chi^2=\sum\frac{(O_i-E_i)^2}{E_i}}$ O_i - שכיחות נצפית של קטגוריה i $\quad p_i$ - הסתברות לקטגוריה i לפי השערת האפס. $E_i=np_i$ שכיחות צפויה של קטגוריה i.

מבחן לאי תלות בין משתנים: מבחן לאי תלות מטרתו לבדוק האם קיים קשר בין שני משתנים: שני המשתנים שנבדקים צריכים להיות מחולקים למספר קטגוריות . למשל, X – מין: זכר ונקבה. Y – דעה פוליטית: ימין , מרכז שמאל.

השערות: H_0: אין תלות בין X ל- Y

H_1: יש תלות בין X ל- Y

כלל הכרעה: נדחה את H_0 אם נקבל כי הערך הסטטיסטי גבוה מהערך הקריטי.

$df=(l-1)(m-1)$ $\quad k=\chi^{2(df)}_{1-\alpha}$ $\quad l$ = מספר הקטגוריות בעמודות (למשל: ימין, מרכז, שמאל- 3) m = מספר הקטגוריות בשורות (למשל: נשים, גברים-2)

הסטטיסטי: $\chi^2=\sum\frac{(O_i-E_i)^2}{E_i}$ O_i - שכיחות נצפית של משבצת i $\quad p_i$ - הסתברות למשבצת i לפי השערת האפס. $E_i=\frac{f_x\cdot f_y}{n}$ שכיחות צפויה של משבצת i.

משתנה רציף:

$\boxed{E(X)=\int X\cdot f(x)dx=\mu}$ $\boxed{V(X)=\int X^2f(x)dx-\mu^2=\sigma^2}$ $\boxed{F(t)=p(x\leq t)}$ $\boxed{p(x>t)=1-F(t)}$ $\boxed{p(a<x<b)=F(a)-F(b)}$

התפלגות מעריכית: התפלגות זו היא התפלגות רציפה המאפיינת את הזמן עד להתרחשות מאורע מסויים. בתרגילים העוסקים בהתפלגות זו יאמר שהמשתנה מתפלג מעריכית:

פונקציית הצפיפות של ההתפלגות היא: $f(x)=\lambda e^{-\lambda x}$ פונקציית ההתפלגות המצטברת היא: $\boxed{F(t)=p(x\leq t)=1-e^{-\lambda t}}$

λ - הוא ממוצע הארועים ביחידת זמן. $E(x)=\frac{1}{\lambda}, V(x)=\frac{1}{\lambda^2}$