

Konfidensintervall för skillnad i väntevärden

Modell: X_1, \dots, X_{n_x} är $N(m_x, \sigma_x)$ och Y_1, \dots, Y_{n_y} är $N(m_y, \sigma_y)$ och alla stokastiska variabler är oberoende. Vi vill sätta upp konfidensintervall för skillnaden i väntevärden $m_x - m_y$. Skillanden skattas med $\bar{x} - \bar{y}$ som beskrivs av

$$\bar{X} - \bar{Y} \text{ är } N\left(m_x - m_y, \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}\right).$$

Vi skiljer mellan tre fall:

1. σ_x och σ_y kända
2. σ_x och σ_y okända men förmodat lika, $\sigma_x = \sigma_y = \sigma$.
3. σ_x och σ_y okända och förmodat olika, $\sigma_x \neq \sigma_y$.

I fall 1 kan vi med en gång sätta upp konfidensintervallet med konfidensgrad $1 - \alpha$:

$$\begin{aligned} m_x - m_y &\in (\bar{X} - \bar{Y}) \pm \lambda_{\alpha/2} D (\bar{X} - \bar{Y}) \\ &\in (\bar{X} - \bar{Y}) \pm \lambda_{\alpha/2} \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}. \end{aligned}$$

I det fall 3 har vi den approximativa metoden:

$$\begin{aligned} m_x - m_y &\in (\bar{X} - \bar{Y}) \pm \lambda_{\alpha/2} D^* (\bar{X} - \bar{Y}) \\ &\in (\bar{X} - \bar{Y}) \pm \lambda_{\alpha/2} \sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}. \end{aligned}$$

är ett approximativt $1 - \alpha$ konfidensintervall för $m_x - m_y$.

Om σ_x och σ_y okända men förmodat lika, $\sigma_x = \sigma_y = \sigma$, är

$$\frac{(\bar{X} - \bar{Y}) - (m_x - m_y)}{\sigma \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \text{ är } N(0, 1)$$

och

$$\frac{(\bar{X} - \bar{Y}) - (m_x - m_y)}{S \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \text{ är } t(\cdot)$$

där antalet frihetsgrader beror på skattningsvariabeln S av σ . Ett $1 - \alpha$ [symmetriskt] konfidensintervall ges av

$$m_x - m_y \in (\bar{X} - \bar{Y}) \pm t_{\alpha/2} S \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}.$$

Den bästa variansskattningen utnyttjar både s_x^2 och s_y^2 : Den sammanvägda (poolade) stickprovsvariansen

$$s^2 = \frac{\sum_{i=1}^{n_x} (x_i - \bar{x})^2 + \sum_{i=1}^{n_y} (y_i - \bar{y})^2}{n_x + n_y - 2} = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}.$$

är en skattning av σ^2 där

$$\frac{(n_x + n_y - 2)S^2}{\sigma^2} = \underbrace{\frac{1}{\sigma^2} \sum_{i=1}^{n_x} (X_i - \bar{X})^2}_{=\chi^2(n_x-1)} + \underbrace{\frac{1}{\sigma^2} \sum_{i=1}^{n_y} (Y_i - \bar{Y})^2}_{=\chi^2(n_y-1)}$$

är χ^2 -fördelad med $n_x + n_y - 2$ frihetsgrader. Alltså är

$$\frac{(\bar{X} - \bar{Y}) - (m_x - m_y)}{S \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \text{ är } t(n_x + n_y - 2)\text{-fördelad.}$$

Exempel: Fritt från TEMO:s väljarbarometer. Vid $n_1 = 2854$ intervjuer (mars) erhöles $x_1 = 1085$ (s)-sympatisörer. I april var motsvarande $x_2 = 1452$ av $n_2 = 3762$ tillfrågade. Andelarna sympatisörer p_1 och p_2 vid de två tillfällena skattas med

$$p_1^* = \frac{x_1}{n_1} = \frac{1085}{2854} = 0.380 \quad p_2^* = \frac{x_2}{n_2} = \frac{1452}{3762} = 0.386$$

vilket ger en skattning av förändringen $p_2 - p_1$: $p_2^* - p_1^* = 0.006$. Konfidensintervall:

$$p_2^* - p_1^* \text{ är approx } N \left(p_2 - p_1, \sqrt{\frac{p_2(1-p_2)}{n_2} + \frac{p_1(1-p_1)}{n_1}} \right)$$

Detta ger konfidensintervallet

$$p_2 - p_1 \in p_2^* - p_1^* \pm \lambda_{\alpha/2} \sqrt{\frac{p_2^*(1-p_2^*)}{n_2} + \frac{p_1^*(1-p_1^*)}{n_1}} = 0.006 \pm 0.024 \quad (\approx 95\%)$$

med konfidensgrad approximativt 95%. Vi ser att $p_2 - p_1 = 0$ inte är ett orimligt värde enligt konfidensintervallet. Vi har ingen statistiskt säkerställd förändring.

Konfidensintervall för skillnad i väntevärden vid parade försök Här är X_1, \dots, X_n och Y_1, \dots, Y_n inte oberoende eller ens nödvändigtvis likafördelade. Detta uppstår typiskt vid mätningar på enheter före (X_i) och efter (Y_i) efter en insats/behandling. Modellen är att *paren*

$$(X_1, Y_1), \dots, (X_n, Y_n)$$

är sinsemellan oberoende, och behandlingseffekterna

$$\Delta_i = Y_i - X_i \text{ är } N(m, \sigma)$$

Notera att vi gör inget uttalande om fördelningarna för X_i och Y_i .

Hypotesprövning

Hypoteser är utsagor om parametervärden θ . Vi ställer vanligtvis upp två hypoteser:

$$H_0 : \theta \in \Theta_0,$$

en *nollhypotes*, och

$$H_1 : \theta \notin \Theta_0,$$

en *mothypotes/alternativhypotes*. Vi vill med observationer X_1, \dots, X_n försöka motbevisa utsagan H_0 . Till vår hjälp har vi en teststatistika

$$T = g(X_1, \dots, X_n)$$

som, då H_0 är sann, har en känd fördelning. (Den ges oftast av en skattningsvariabel av θ). Då kan vi bestämma en mängd rimliga värden för T , dvs. värden på T som är förenliga med H_0 , och en mängd orimliga värden, betecknad C . Mängden C kallas *förkastelseområde* eller *kritiskt område*.

Beslutsregel: Om vi observerar ett utfall $T \in C$ så förkastar vi H_0 , annars inte. Vi kan tänka oss fyra scenarion:

	H_0 sann	H_0 falsk
$T \notin C$	OK.	Fel av 2:a slaget
$T \in C$	Fel av 1:a slaget	OK.

Vi vill begränsa sannolikheten för fel av första slaget, dvs sannolikheten att förkasta en korrekt nollhypotes. Vi väljer C så att

$$P(T \in C | H_0 \text{ sann}) \leq \alpha$$

för ett litet α , kallat *signifikansnivån*. Funktionen

$$\beta(\theta) = P(T \in C) = P(\text{Förkasta } H_0)$$

kallas testets styrkefunktion.

Schematisk beskrivning av ett hypotesprövningsförfarande

1. Formulera hypoteser och fixera en signifikansnivå α .
2. Tag fram en teststatistika T som skall skilja hypoteserna åt, och har en känd fördelning om H_0 är sann.
3. Bestäm förkastelseområdet C , värden orimliga under H_0 men mer troliga då mothypotesen sann.
4. Gör mätningar och observera utfallet på teststatistikan.

Exempel: Opinionsundersökning: test ifall andelen sympatisörer har förändrats Vill testa

$$H_0 : p_1 = p_2 \text{ mot } H_1 : p_1 \neq p_2$$

på signifikansnivå approximativt $\alpha = 0.05$. Hypoteserna formuleras

$$H_0 : p_2 - p_1 = 0 \text{ mot } H_1 : p_2 - p_1 \neq 0.$$

2. Skatta $p_2 - p_1$ med $p_2^* - p_1^*$. Med

$$T = \frac{p_2^* - p_1^*}{\sqrt{\frac{p_2^*(1-p_2^*)}{n_2} + \frac{p_1^*(1-p_1^*)}{n_1}}}$$

så är T approximativt $N(0, 1)$ om H_0 är sann.

3. Förkasta H_0 då $|T|$ är stor. Ur $N(0, 1)$ -tabeller fås att $P(|T| > \lambda_{\alpha/2}) = \alpha$. Med $\alpha = 0.05$ fås $\lambda_{0.025} = 1.96$. Alltså: förkasta H_0 om $|T| > 1.96$.

4. Vid $n_1 = 2854$ intervjuer (mars) erhöles $x_1 = 1085$ (s)-sympatisörer. I april var motsvarande $x_2 = 1452$ av $n_2 = 3762$ tillfrågade.

$$p_1^* = \frac{x_1}{n_1} = \frac{1085}{2854} = 0.380 \quad p_2^* = \frac{x_2}{n_2} = \frac{1452}{3762} = 0.386$$

och

$$t = \frac{p_2^* - p_1^*}{\sqrt{\frac{p_2^*(1-p_2^*)}{n_2} + \frac{p_1^*(1-p_1^*)}{n_1}}} = 0.497 < 1.96$$

och vi förkastar inte H_0 till förmån för H_1 på nivå 5%.