

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 de parvisa skillnaderna

$$w_i = y_i - x_i, \quad i = 1, 2, \dots, n$$

är utfall av oberoende $N(\mu, \sigma)$ -fördelade stokastiska variabler $W_1, \dots, W_n$. Notera att här är

$$Y_i = X_i + W_i$$

där inga modelleringsantaganden ligger på fördelningarna för X_i och Y_i .

Hypotesprövning

Hypoteser är vanligtvis utsagor om parametervärden θ . Vi ställer 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 teststorhet

$$t = t(x_1, \dots, x_n)$$

som, då H_0 är sann, är ett utfall på en stokastisk variabel $T = t(X_1, \dots, X_n)$ med känd fördelning. (Den ges oftast av en skattning 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 $t \in C$ 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 *signifikansnivå*.

Funktionen

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

kallas testets styrkefunktion.

Exempel: Fortkörning. Låt x vara mätningen av en bilists hastighet. Modell x är ett utfall av X där

$$X = \text{hastighet} + \text{mätfel} = v + \epsilon, \quad \epsilon \in N(0, \sigma),$$

dvs X är $N(v, \sigma)$. Antag att $\sigma = 1.2$ km/h.

Vill testa ifall en bilist har en hastighet större hastighet än $v_0 = 70$ km/h.

Vill testa

$$H_0 : v = v_0 \text{ mot } H_1 : v > v_0$$

på signifikansnivå approximativt $\alpha = 0.01$.

Vi förkastar H_0 till förmån för H_1 för stora värden på

$$t = \frac{x - v_0}{\sigma} = \frac{x - 70}{1.2}$$

som om H_0 är sann är ett utfall av en $N(0, 1)$ -fördelad stokastisk variabel T .

Ur $N(0, 1)$ -tabeller fås att $P(T > \lambda_\alpha) = \alpha$. Med $\alpha = 0.01$ fås $\lambda_{0.01} = 2.3263$. Alltså: förkasta H_0 om $t > 2.3263$.

Om man observerar utfallet $x = 72.7$ km/h så är

$$t = \frac{72.7 - 70}{1.2} = 2.25 < 2.3262$$

och vi förkastar ej H_0 på nivå 1%.

Den lägsta signifikansnivå som man kan förkasta observationen t på kallas för testets p-värde. Här

$$\text{p-värdet} = P(T > 2.25) = 1 - \Phi(2.25) = 0.01222$$

så vi förkastar t.ex. H_0 på nivå 5%.

Notera att vi förkastar ej H_0 då $(x - v_0)/\sigma \leq \lambda_\alpha$ dvs då $v_0 \geq x - \sigma\lambda_\alpha$; ett ensidigt konfidensintervall för hastigheten (konfidensgrad $1 - \alpha$). Här fås $v_0 \geq 72.7 - 1.2 \cdot 2.3263 = 69.91$ km/h (99%).

χ^2 -test

Observationer kategoriseras i r kategorier. Antag att en observation har sannolikhet p_i att hamna i kategori i , $i = 1, \dots, r$. $\sum_{i=1}^r p_i = 1$. Av n oberoende försök låt X_i beskriva antalet observationer i kategori i . Modell:

$$X_i \sim \text{Bin}(n, p_i), \quad i = 1, \dots, r.$$

Det förväntade antalet observationer i kategori i är

$$E[X_i] = np_i.$$

Notera att $X_1, \dots, X_r$ inte är oberoende. $\sum_{i=1}^r X_i = n$. Vektorn $(X_1, \dots, X_r)$ är multinomialfördelad

$$P(X_1 = x_1, \dots, X_r = x_r) = \frac{n!}{x_1! \cdots x_r!} p_1^{x_1} \cdots p_r^{x_r}.$$

Med $r = 2$ fås binomialfördelningen.

Sats.

$$Q = \sum_{i=1}^r \frac{(X_i - np_i)^2}{np_i} \underset{\text{approx}}{\sim} \chi^2(r-1)$$

Alla test framgent bygger på detta resultat. Approximationen fungerar bäst då $np_i \geq 5$, dvs. det förväntade antalet observationer i varje kategori får inte vara för litet.

Exempel (Test på andelar/fördelning): Låt X beskriva antalet bilar på en parkeringsplats. Vill testa

$$H_0 : X \in \text{Po}(4)$$

mot

$$H_1 : X \notin \text{Po}(4)$$

på nivå 5%.

Om H_0 är sann så är

$$\begin{aligned} p_1 &= P(X \leq 2) = 0.2381 \\ p_2 &= P(X = 3) = 0.1954 \\ p_3 &= P(X = 4) = 0.1954 \\ p_4 &= P(X \geq 5) = 0.3712 \end{aligned}$$

50 observationer sammanfattas av

	kategori (antal bilar)				
	1 (0-2)	2 (3)	3 (4)	4 (5-)	
Observerat antal: x_i	22	15	6	7	50
Förväntat antal: np_i	11.9	9.8	9.8	18.6	50

Vi förkastar hypotesen för stora värden på

$$q = \sum_i \frac{(x_i - np_i)^2}{np_i} = 20.014$$

som om H_0 är sann är ett utfall på en $\chi^2(3)$ -fördelad stokastisk variabel. Ur tabell fås $\chi_\alpha^2 = 7.8147$. Då $q > \chi_\alpha^2$ förkastas H_0 på nivå $\alpha = 5\%$. Det är orimligt att antalet bilar X är Poisson(4)-fördelad.

För att testa hypoteser som $H_0 : X$ är $\text{Po}(\mu)$, något μ . Skatta μ med $\mu_{\text{obs}}^* = 2.82$. Skattningen ger

	kategori (antal bilar)				
	1 (0-2)	2 (3)	3 (4)	4 (5-)	
Observerat antal: x_i	22	15	6	7	50
Hypotes: $p_i(\mu_{\text{obs}}^*)$	0.465	0.223	0.157	0.155	1.0
Förväntat antal: $np_i(\mu_{\text{obs}}^*)$	23.2	11.1	7.9	7.8	50

Vi förkastar hypotesen för stora värden på

$$q = \sum_i \frac{(x_i - np_i(\mu_{\text{obs}}^*))^2}{np_i(\mu_{\text{obs}}^*)} = 1.92$$

som om H_0 är sann är ett utfall på en $\chi^2(4 - 1 - 1)$ -fördelad stokastisk variabel. Ur tabell fås $\chi_\alpha^2 = 5.99$. Då $q < \chi_\alpha^2$ förkastas inte H_0 på nivå $\alpha = 5\%$. Det är inte orimligt att antalet bilar X är Poisson-fördelad.

Generellt,

$$(\text{Antalet frihetsgrader}) = (\text{Antalet kategorier} - 1) - (\text{Antalet skattade parameterar}).$$