

Hypotesprövning

Vi ska idag ta fram metoder för hypotesprövning, dvs allt med matematiska metoder avgöra om en uppställd hypotes stöds av de givna observationerna.

Ex: En person säger sig kunna påverka utgången vid slantsingling så att $p = IP(\text{krona}) > 1/2$. För att testa om personen har fög för sitt påstående ställer vi upp en nollhypotes H_0 som lyder

$H_0: p = 1/2$ (personen kan inte påverka utgången vid slantsinglingen)

Om personen vid ett antal slantsinglingar får krona tillräckligt många gånger så är vi beredda att tro hans påstående, dvs att förkasta nollhypotesen. Vi antar nu att personen utför 10 slantsinglingar och låter x beteckna antalet gånger som utfallet är krona. x är en observation av $X \in \text{Bin}(10, p)$. Om H_0 är sann, så gäller $X \in \text{Bin}(10, 1/2)$. Vårt hypotestest kommer att ha formen

"Förkasta H_0 om $x \geq a$ för lämpligt val av konstanten a ".

Vi vill välja a så att vi får så liten felrisk (eller signifikansnivå) som möjligt, dvs risken att förkasta H_0 om H_0 är sann (att påstå att personen kan påverka slantsinglingen när han inte kan det) ska vara liten. Vi undersöker $IP(X \geq a)$ för $X \in \text{Bin}(10, 1/2)$ och olika a .

$$IP(X \geq a) = \sum_{k=a}^{10} \binom{10}{k} \left(\frac{1}{2}\right)^{10} \text{ ger tabellen}$$

a	10	9	8	7
$IP(X \geq a)$	$9,8 \cdot 10^{-3}$	0,011	0,055	0,172

Om felrisken ska vara högst 5%, bör vi välja $a=9$, dvs förkasta H_0 ($p=1/2$) om det blir krona minst nio gånger. Felrisken blir då 1,1%.

Vi formulerar nu hypotesprövning allmänt. Låt $x = (x_1, \dots, x_n)$ vara ett stickprov som är ett utfall av en sv. $X = (X_1, \dots, X_n)$ vars fördelning beror av en okänd parameter θ . Vi vill testa en grundhypotes eller nollhypotes H_0 om θ , ex. vis $\theta = \theta_0$, mot en alternativ hypotes eller mothypotes H_1 om θ . Mothypotesen kan vara av följande slag:

Vi antar att H_0 är sann och med hjälp av informationen i x kommer vi ansträngen att förkasta H_0 eller inte kunna förkasta H_0 . Notera att "förkasta H_0 " inte med säkerhet innebär att H_1 är sann och att "inte kunna förkasta H_0 " inte med säkerhet innebär att H_0 är sann. För att testa H_0 definierar vi en testvariabel $t_{\text{obs}} = t(x)$ som är en observation av motsvarande stickprövsvariabel $t(X)$ (i exemplet var $t(x) = x$). Till testvariabeln bestämmer vi ett kritiskt område C och vårt hypotestest utformas som

Det kritiska området väljs så att

$$IP(H_0 \text{ förkastas} \mid H_0 \text{ sann}) = IP(t(X) \in C) = \alpha \text{ för någon förbestämd felrisk } \alpha.$$

Vanliga värden på α är 5%, 1% och 0,1%. I exemplet var $C = \{9, 10\}$ för felrisken 5%. Om vi hade fått $x=7$, så hade $x \notin C$ och vi hade inte kunnat förkasta H_0 (dvs vi har inte tillräckliga belegg för att personen kan påverka utgången av slantsinglingen).

Vi arbetar ofta med flera signifikansnivåer samtidigt. Med signifikant* menas att ett test har felrisk mellan 1% och 5% (signifikant på 5%-nivån men inte på 1%-nivån). Signifikant** och signifikant*** definieras analogt (se boken).

I stället för att bestämma ett kritiskt område C för testvariabeln t , så kan vi basera ett hypotestest på det så kallade p -värdet, som är ett mått på hur orimlig H_0 är givet stickprovet x . Om p -värdet är mindre än den givna felrisken α , så förkastar vi H_0 .

Def: p -värdet definieras som $IP(t(X))$ är minst lika extremt som $t(x) | H_0$ sann

I slantsinglingsexemplet är p -värdet för $x=9$ lika med $IP(X \geq 9) = 0,011$, så på signifikansnivån 5%, så förkastar vi H_0 om minst nio av de tre slantsinglingarna resulterar i krona.

Ex: I riksdagsvalet 2014 fick Centerpartiet 6,1% av rösterna, medan i SCB:s partisympatibundersökning från i november i år svarade 357 av $n=5021$ svarande (7,1%) att de skulle rösta på Centerpartiet om det var val nu. Vi vill testa om andelen C -sympatisörer p i populationen ökat sedan valet. Låt H_0 : "Oförändrad andel" och H_1 : "Ökad andel". Om H_0 är sann så är antalet $\bar{X}$ som i SCB-undersökningen svarar att de skulle rösta på Centerpartiet $\bar{X} \sim \text{Hyp}(N, n, p) \approx \text{Bin}(n, p) \approx N(np, \sqrt{np(1-p)}) \approx N(306, 16,96)$

Antalet röstberättigade $N \gg n$ $np(1-p) = 288 \gg 10$
 p -värdet för det observerade antalet C -väljare $x=357$ är:

$$IP(\bar{X} \geq 357) = IP\left(\frac{\bar{X} - np}{\sqrt{np(1-p)}} \geq \frac{357 - np}{\sqrt{np(1-p)}}\right) = 1 - \Phi\left(\frac{357 - 306}{16,96}\right) = 0,0013$$

så vi förkastar H_0 på signifikansnivån 0,0013. Antalet C -sympatisörer har ökat och resultatet är signifikant**.

Vi kan också använda konfidensintervall för att utföra hypotesprövning. För nollhypotesen $H_0: \theta = \theta_0$ och felrisken α , så bestämmer vi i så fall ett konfidensintervall I_α för θ med konfidensgrad $1-\alpha$. Om nollhypotesen är tvåsidig, dvs $H_1: \theta \neq \theta_0$, så väljer vi ett tvåsidigt konfidensintervall, men om nollhypotesen är ensidig så ska konfidensintervallet vara ensidigt ($H_1: \theta > \theta_0$ motsvarar $I_\alpha = (\alpha, \infty)$). Vi förkastar H_0 om $\theta_0 \notin I_\alpha$ eftersom då gäller

$$IP(H_0 \text{ förkastas} | H_0 \text{ sann}) = IP(\theta_0 \notin I_\alpha) = 1 - IP(\theta_0 \in I_\alpha) = 1 - (1 - \alpha) = \alpha$$

Ex: Låt $x = (x_1, \dots, x_n)$ vara ett stickprov av en s.v. $\bar{X} = (\bar{X}_1, \dots, \bar{X}_n)$, där $\bar{X}_i \sim N(\mu, \sigma)$ för okända μ och σ . Vi vill testa nollhypotesen $H_0: \mu = \mu_0$ mot nollhypotesen $H_1: \mu > \mu_0$. Eftersom nollhypotesen är ensidig så bestämmer vi ett ensidigt konfidensintervall för μ med konfidensgrad $1-\alpha$. En lämplig testvariabel är

$$t(\bar{X}) = \frac{\bar{X} - \mu}{S/\sqrt{n}} \in t(n-1)$$

och med hjälp av t -fördelningens kvantiler fås:

$$IP(t(\bar{X}) < t_\alpha(n-1)) = 1 - \alpha \Rightarrow IP(\bar{X} - \mu < t_\alpha(n-1) \frac{S}{\sqrt{n}}) = 1 - \alpha$$

$$\Rightarrow IP(\bar{X} - t_\alpha(n-1) \frac{S}{\sqrt{n}} < \mu) = 1 - \alpha \Rightarrow I_\mu = (\bar{X} - t_\alpha(n-1) \frac{S}{\sqrt{n}}, \infty)$$

Vi förkastar nollhypotesen om $\mu_0 \notin I_\mu$, dvs om $\mu_0 < \bar{X} - t_\alpha(n-1) \frac{S}{\sqrt{n}}$ eller, ekvivalent, om $\bar{X} > \mu_0 + t_\alpha(n-1) \frac{S}{\sqrt{n}}$.

Om nollhypotesen istället hade varit $H_1: \mu \neq \mu_0$ så hade vi använt ett tvåsidigt konfidensintervall för μ , dvs

$$I_\mu = (\bar{X} - t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}})$$

Vi hade då förkastat nollhypotesen om $\mu_0 > \bar{X} + t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}}$ eller om $\mu_0 < \bar{X} - t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}}$, dvs om $|\bar{X} - \mu_0| > t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}}$

Notera att de två metoderna alltid ger samma slutsats om H_0 .