Föreläsning 12

Thomas Önskog

6/12 2017

Hypotesprövning

Vi ska idag ta fram metoder för **hypotesprövning**, dvs att med matematiska metoder avgöra om en uppställd hypotes stöds av de givna observationerna eller ej. Vi beskriver nu den generella metodiken och illustrerar parallellt med ett exempel.

Låt $x = (x_1, ..., x_n)$ vara ett stickprov som är ett utfall av en s.v. $X = (X_1, ..., X_n)$ vars fördelning beror av en okänd parameter θ. Vi vill testa en grundhypotes eller **nollhypotes** H_0 om värdet på parametern θ, exempelvis $\theta = \theta_0$, mot en alternativ hypotes eller **mothypotes** H_1 om θ:s värde. Mothypotesen kan vara av följande slag:

$$\begin{array}{ll} H_1 : \theta = \theta_1 & \text{enkel} \\ H_1 : \theta > \theta_0 & \text{sammansatt, ensidig} \\ H_1 : \theta < \theta_0 & \text{sammansatt, ensidig} \\ H_1 : \theta \neq \theta_0 & \text{sammansatt, tvåsidig} \end{array}$$

Vi antar att nollhypotesen H_0 är sann och med hjälp av informationen i stickprovet x kommer vi antingen att kunna **förkasta** H_0 eller **inte kunna förkasta** H_0. Notera att "förkasta H_0" inte med säkerhet innebär att H_0 är falsk och att "inte kunna förkasta H_0" inte med säkerhet innebär att H_0 är sann.

Exempel. *En person säger sig kunna påverka utgången vid slantsingling så att det gäller att $p = \mathbb{P}(\text{krona}) > 1/2$. För att testa om personen har fog för sitt påstående ställer vi upp en nollhypotes H_0 som lyder*

$$H_0 : p = 1/2 \qquad \textit{(personen kan inte påverka utgången vid slantsingling).}$$

och vår mothypotes blir $H_1 : p > 1/2$ (personen kan påverka utgången vid slantsingling så att det oftare blir krona än klave).

För att testa nollhypotesen H_0 definierar vi en **testvariabel** $t_{\text{obs}} = t(x)$ som är en observation av motsvarande stickprovsvariabel $t(X)$. Till testvariabeln bestämmer vi ett **kritiskt område** C och vårt hypotestest utformas som

$$\begin{cases} \text{"Om } t_{\text{obs}} \in C, \text{ så förkastas } H_0.\text{"} \\ \text{"Om } t_{\text{obs}} \notin C, \text{ så förkastas inte } H_0.\text{"} \end{cases}$$

Det kritiska området väljs så att

$$\mathbb{P}(H_0 \text{ förkastas}|H_0 \text{ är sann}) = \mathbb{P}(t(X) \in C|H_0 \text{ är sann}) = \alpha,$$

för någon förutbestämd **felrisk** (eller **signifikansnivå**) α. Vanliga värden på α är 5%, 1% och 0.1%.

Exempel. *Om personen vid ett antal slantsinglingar får krona tillräckligt många gånger, så är vi beredda att tro hens påstående, dvs att förkasta nollhypotesen. Vi antar nu att personen utför* 20 *slantsinglingar och låter* x *beteckna antalet gånger som utfallet är krona.* x *är en observation av* $X \in Bin(20, p)$. *Vi antar att nollhypotesen är sann, så* $X \in Bin(20, 1/2)$. *Vår testvariabel är i detta fall* $t(x) = x$ *och det kritiska området har formen* $C = \{a, ..., 20\}$ *för något heltal* a. *Det kritiska området ska uppfylla*

$$\alpha = \mathbb{P}(t(X) \in C) = \mathbb{P}(X \geq a),$$

så vi undersöker $\mathbb{P}(X \geq a)$ *för några olika värden på* a. *Med hjälp av sannolikhetsfunktionen för binomialfördelningen,* $p_X(k) = \binom{20}{k}\left(\frac{1}{2}\right)^{20}$, *får vi*

a	20	19	18	17	16	15	14
$\mathbb{P}(X \geq a)$	$9.5 \cdot 10^{-7}$	$2.0 \cdot 10^{-5}$	$2.0 \cdot 10^{-4}$	0.0013	0.0059	0.0207	0.0577

Om felrisken ska vara högst 5%, *så bör vi välja* $a = 15$, *dvs förkasta nollhypotesen om personen får krona minst* 15 *gånger av* 20. *Den faktiska felrisken blir då* 2.1%.

Formen på det kritiska området beror på valet av mothypotes. Om mothypotesen i slantsinglingsexemplet hade varit $H_1 : p \neq 1/2$, så hade det kritiska området haft formen $C = \{0, ..., a\} \cup \{20 - a, ..., 20\}$, dvs vi hade förkastat nollhypotesen både om det hade blivit oväntat många krona och om det hade blivit oväntat många klave.

Vi arbetar ofta med flera olika signifikansnivåer samtidigt. Med signifikant* menas att ett test har felrisk mellan 1% och 5%, dvs. resultatet är signifikant på 5%-nivån, men inte på 1%-nivån, såsom i ovanstående exempel. På liknande sätt innebär signifikant** att ett test har en felrisk mellan 0.1% och 1% och signifikant*** att ett test har en felrisk under 0.1%.

Istället 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 orimligt stickprovet x är givet nollhypotesen H_0. Om p-värdet är mindre än den givna felrisken α, så förkastar vi H_0.

Definition. ***p-värdet*** *definieras som* $\mathbb{P}(t(X)$ *är minst lika extremt som* $t(x)|H_0$ *är sann).*

Exempel. *I slantsinglingsexemplet är p-värdet för* $x = 15$ *lika med* $\mathbb{P}(X \geq 15) = 0.0207 \leq 0.05$, *så på signifikansnivån* 5% *så förkastar vi* H_0 *om minst* 15 *av de* 20 *slantsinglingarna resulterar i krona.*

Precis som för det kritiska området, så beror p-värdet på valet av mothypotes. Om mothypotesen hade varit $H_1 : p \neq 1/2$, *så hade p-värdet för* $x = 15$, *varit* $\mathbb{P}(X \geq 15) + \mathbb{P}(X \leq 5) = 0.0207 + 0.0207 = 0.0414$. *Även i detta fall hade vi dock förkastat* H_0 *på signifikansnivån* 5%.

Vi tar ytterligare ett exempel på hypotestest med hjälp av p-värde.

Exempel. *I riksdagsvalet* 2014 *fick Moderaterna* 23.3% *av rösterna. I SCB:s partisympatiundersökning från i november i år svarade* 1047 *av* $n = 4715$ *svarande (*22.2%*) att de skulle rösta på Moderaterna om det var val nu. Vi vill testa om andelen M-sympatisörer* p *i populationen minskat sedan valet* 2014*. Låt*

$$H_0 : p = 0.233 \text{ (Oförändrad andel M-sympatisörer)}$$

$$H_1 : p < 0.233 \text{ (Minskad andel M-sympatisörer)}$$

Om H_0 *är sann, så är antalet personer som i SCB-undersökningen svarade att de skulle rösta på M ett utfall av en* $Hyp(N, n, p)$*-fördelad s.v.* X*, där* N *är antalet personer som är röstberättigade i det svenska riksdagsvalet,* $n = 4715$ *är antalet svarande och* $p = 0.233$ *är andelen M-sympatisörer under nollhypotesen. Eftersom* $N \gg n$*, så är* $X \approx Bin(n, p)$ *och eftersom* $np(1-p) = 843 \geq 10$*, så är* $X \approx N(np, \sqrt{np(1-p)})$*.* p*-värdet för det observerade antalet M-väljare är därmed*

$$\mathbb{P}(X \leq 1047) = \mathbb{P}\Big(\underbrace{\frac{X - np}{\sqrt{np(1-p)}}}_{\in N(0,1)} \leq \frac{1047 - np}{\sqrt{np(1-p)}}\Big) = \Phi\Big(\frac{1047 - 1098.6}{29.03}\Big) = \Phi(-1.78) = 0.038,$$

så vi förkastar H_0 *på signifikansnivån* 0.038*. Antalet M-sympatisörer har minskat 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_θ för θ med konfidensgrad $1 - \alpha$. Om mothypotesen är tvåsidig, dvs $\theta \neq \theta_0$, så väljer vi ett tvåsidigt konfidensintervall, men om mothypotesen är ensidig, så ska konfidensintervallet vara ensidigt ($H_1 : \theta > \theta_0$ motsvarar $I_\theta = (a, \infty)$). Vi förkastar H_0 om $\theta_0 \notin I_\theta$, eftersom det gäller att

$$\begin{aligned}
\mathbb{P}(H_0 \text{ förkastas} | H_0 \text{ är sann}) &= \mathbb{P}(\theta_0 \notin I_\theta | H_0 \text{ är sann}) = 1 - \mathbb{P}(\theta_0 \in I_\theta | H_0 \text{ är sann}) \\
&= 1 - (1 - \alpha) = \alpha.
\end{aligned}$$

Exempel. *Låt* $x = (x_1, ..., x_n)$ *vara ett stickprov som är ett utfall av en s.v.* $X = (X_1, ..., X_n)$*, där* $X_i \in N(\mu, \sigma)$ *för okända* μ *och* σ*. Vi vill testa nollhypotesen* $H_0 : \mu = \mu_0$ *mot mothypotesen* $H_1 : \mu > \mu_0$*. Eftersom mothypotesen är ensidig, så bestämmer vi ett ensidigt konfidensintervall för* μ *med konfidensgrad* $1 - \alpha$*. En lämplig testvariabel är*

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

och med hjälp av t-fördelningens kvantiler fås

$$\begin{aligned}
\mathbb{P}\Big(\frac{\overline{X} - \mu}{S/\sqrt{n}} < t_\alpha(n-1)\Big) = 1 - \alpha \qquad &\Rightarrow \qquad \mathbb{P}\Big(\overline{X} - \mu < t_\alpha(n-1)\frac{S}{\sqrt{n}}\Big) = 1 - \alpha \\
&\Rightarrow \qquad \mathbb{P}\Big(\overline{X} - t_\alpha(n-1)\frac{S}{\sqrt{n}} < \mu\Big) = 1 - \alpha \\
&\Rightarrow \qquad I_\mu = \Big(\overline{x} - t_\alpha(n-1)\frac{s}{\sqrt{n}}, \infty\Big).
\end{aligned}$$

Vi förkastar nollhypotesen H_0 *om* $\mu_0 \notin I_\mu$, *dvs om* $\mu_0 < \overline{x} - t_\alpha(n-1)s/\sqrt{n}$. *Med andra ord förkastar vi nollhypotesen om*

$$\overline{x} > \mu_0 + t_\alpha(n-1)\frac{s}{\sqrt{n}},$$

dvs. om punktskattningen $\overline{x}$ *av* μ *blir tillräckligt mycket större än det värdet* μ_0 *som nollhypotesen anger.*

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

$$I_\mu = \left(\overline{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}, \overline{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\right).$$

Vi hade i detta fall förkastat nollhypotesen om $\mu_0 > \overline{x} + t_{\alpha/2}(n-1)s/\sqrt{n}$ *eller om* $\mu_0 < \overline{x} - t_{\alpha/2}(n-1)s/\sqrt{n}$. *Med andra ord hade vi förkastat nollhypotesen om*

$$|\overline{x} - \mu_0| > t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}},$$

dvs. om punktskattningen $\overline{x}$ *av* μ *avviker tillräckligt mycket (i positiv eller negativ riktning) från det värdet* μ_0 *som nollhypotesen anger.*

Notera att de tre metoderna för hypotestester som vi har beskrivit idag alltid ger samma slutsats om nollhypotesen H_0.