SF1901

HYPOTESPRÖVNING (FÖRELÄSNING 22/2)

Vi betraktar ett stickprov $\mathbf{x} = (x_1, \dots, x_n)$ från en stokastisk variabel $\mathbf{X}$. Inledningsvis antar vi att fördelningen för $\mathbf{X}$ innehåller en okänd parameter θ - vi kommer även se hypotesprövningar som inte rör en specifik parameter.

Målet med hypotesprövning, även kallat ett statistiskt test, är att betrakta en grundhypotes, *nollhypotesen*, H_0 jämfört med en alternativ hypotes H_1 givet data. Följande utgör en form av statistiskt test (*signifikanstest*):

- En nollhypotes H_0 och en alternativ hypotes H_1.
- En *signifikansnivå* $\alpha \in (0, 1)$.
- En *testvariabel* $t(\mathbf{x})$; motsvarande stickprovsvariabel $t(\mathbf{X})$.
- Ett *kritiskt omr{aade* C (en delmn̈gd) av alla möjliga värden för $t(\mathbf{x})$.

Ett signifikantstest kan sammanfattas med fl̈jande beslutsregel:

$$
\begin{aligned}
t(\mathbf{x}) \in C &\Rightarrow \text{ Förkasta } H_0, \\
t(\mathbf{x}) \notin C &\Rightarrow \text{ Förkasta ej } H_0.
\end{aligned}
$$

Det kritiska området C är valt så att, för den valda signifikansnivån α,

$$P(t(\mathbf{X}) \in C) \leq \alpha, \quad \text{om } H_0 \text{ sann.}$$

I fallet med kontinuerliga s.v. kan olikheten (oftast) ersättas med likhet, dvs $P(t(\mathbf{X}) \in C) = \alpha$ om H_0 sann. Signifikansnivån α bör vara liten - vi "skyddar" oss då mot att förkasta H_0 om H_0 är sann (s.k. typ-I-fel).

Hypoteserna H_0 och H_1 kan vara av flera slag. Den för oss främst förekommande typen är s.k. *enkel* hypotes:

$$H_0 : \quad \theta = \theta_0,$$

för något bestämt θ_0. För mothypotesen / alternativa hypotesen H_1 stöter vi på enkla, sammanfattade ensidiga, t.ex.

$$H_1 : \quad \theta > \theta_0,$$

och sammanfattade tvåsidiga,

$$H_1 : \quad \theta \neq \theta_0.$$

Beroende på valet av H_1 kommer det kritiska området att ta olika form. Det illustreras enklast via ett exempel.

Example 0.1. *Låt* $\mathbf{x}$ *vara ett stickprov frå* $N(\mu, \sigma)$, *med* μ, σ *okända. Vi är intresserade av att testa nolllhypotesen*

$$H_0: \ \mu = \mu_0,$$

för ett givet μ_0 *(det värde vi, av någon anledning tror är det "sanna") mot alternativet*

$$H_1: \ \mu \neq \mu_0.$$

Eftersom vi är intresserade av medelvärdet μ verkar det naturligt att betrakta stickprovsmedelvärdet $\bar{x}$. *Likt konstruktionen av konfidensintervall, då vi vill räkna på sannolikheter för stickprovsvariabeln* $t(\mathbf{X})$ *inser vi snart att ett bättre val av testvariabel är*

$$t(\mathbf{x}) = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}.$$

Varför är μ_0 *med? För att under* H_0 *så gäller* $t(\mathbf{X}) \in t(n-1)$ *; vid känd varians* σ^2 *hade vi valt som testvariabel den pivotvariabel som används för konfidensintervall med känd* σ.

Antag att vi vill göra testet på nivån α. *Vi vill nu välja ett område* C *sådant att under* H_0, *markerat med* P_{μ_0}, *gäller*

$$P_{\mu_0}(t(\mathbf{X}) \in C) = \alpha.$$

Hur ska vi välja C*? Om* H_1 *gäller, snarare än* H_0, *så kommer typiskt avståndet* $|\bar{x} - \mu_0|$ *att vara stort.* [1] *Det leder till att vi vill välja en under gräns* l *och en övre gräns* u *s.a.*

$$P_{\mu_0}(\{t(\mathbf{X}) < l\} \cup \{t(\mathbf{X}) > u\}) = \alpha.$$

Eftersom $t(\mathbf{X}) \in t(n-1)$ *kan vi välja* $\alpha/2$*-kvantilen* $t_{\alpha/2}(n-1)$ *från den relevanta t-fördelningen. Dvs., sätt* $l = -t_{\alpha/2}(n-1)$ *och* $u = t_{\alpha/2}(n-1)$, *vilket resulterar i valet*

$$C = (-\infty, -t_{\alpha/2}(n-1)) \cup (t_{\alpha/2}(n-1), \infty).$$

Vi kan även välja att representera det kritiska området via dess komplement (uppenbart varför snart):

$$C^c = (-t_{\alpha/2}(n-1), t_{\alpha/2}(n-1)).$$

Vi förkastar inte H_0 *om* $t(\mathbf{x})$ *hamnar i* C^c. *I det här exemplet förkastar vi alltså inte* H_0 *om*

$$-t_{\alpha/2}(n-1) < t(\mathbf{x}) < t_{\alpha/2}(n-1)),$$

och en omskrivning som i fallet med konfidensintervall visar att det är ekvivalent med

$$\bar{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} < \mu_0 < \bar{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}.$$

Vi kan alltså forma intervallet $I = \bar{x} \pm t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}$ *och se om det innehåller det föreslagna värdet* μ_0 *(förkasta ej* H_0*) eller ej (förkasta* H_0*). Detta visar att ett hypotesttest på nivån* α *kan genomföras genom att skapa ett konfidensintervall med konfidensgrad* $1 - \alpha$ *för parametern av intresse.*

---

[1] Varför? Tänk på fallet då n är stort och att det sanna värdet på μ avviker betydligt från μ_0.

Example 0.2. *Låt allt vara som i det tidigare exemplet, med undantaget att vi nu betraktar den ensidiga mothypotesen*

$$H_1: \ \mu > \mu_0.$$

Idén och tillvägagångssättet "ar detsamma, men nu är det istället så att om H_1 gäller, snarare än H_0, så kommer testvariabeln $t(\mathbf{x})$ typiskt vara stor ($\bar{x}$ större än $\mu_0 \Rightarrow t(\mathbf{x})$ större än 0). Vi väljer därför C enligt

$$C = (t_\alpha(n-1), \infty),$$

ty då fås

$$P_{\mu_0}(t(\mathbf{X}) \in C) = P_{\mu_0}\left(\frac{\bar{x} - \mu_0}{s/\sqrt{n}} > t_\alpha(n-1)\right) = \alpha.$$

På samma sätt som tidigare kan vi skriva om olikheten och ser att vi inte förkastar H_0 då

$$\mu_0 < \bar{x} - t_\alpha(n-1)\frac{s}{\sqrt{n}}.$$

Det senare är ekvivalent med att skapa ett ensidigt konfidensintervall med konfidensgrad 95% för μ och kontrollera huruvida det innehåller μ_0 eller ej.

Exemplen visar hur hypotestest och konfidensintervall kan ses som inverser av varandra: Ett test på nivån 5% svarar mot ett konfidensintervall med konfidensgrad 95% för parametern i fråga. Notera specifikt hur valet av mothypotes H_1 bestämmer huruvida vi ska konstruera ett ensidigt eller tvåsidigt intervall.

Avslutningsvis kan vi betrakta sammansatta nollhypoteser, t.ex.

$$H_0: \quad \theta \leq \theta_0.$$

Det leder oss till att definiera den s.k. *styrkefunktionen* för ett test. Om P_θ betecknar sannolikheter under antagandet att det sanna värdet är θ så definieras styrkefuntionen $h(\theta)$ som

$$h(\theta) = P_\theta(H_0 \text{ förkastas}).$$

För $\theta \in H_0$ (dvs ett värde som ingår i nollhypotesen) gäller då $h(\theta) \leq \alpha$. Vidare så önskar vi att $h(\theta)$ är stor (dvs. nära 1) för $\theta \in H_1$. Omvänt så säger vi att för en sammansatt H_0 så ges signifikansnivån av ett test - kritiskt område C alltså givet - av

$$\alpha = \max_{\theta \in H_0} h(\theta).$$

I ord: signifikansnivån bestäms av det "värsta" möljiga $\theta \in H_0$.