Föreläsning 10: Jämförelser

Matematisk statistik

David Bolin
Chalmers University of Technology
September 27, 2017

Hypotesprövning

Vi har ett stickprov från en fördelning med en okänd parameter θ.

Vi vill testa *nollhypotesen*

$$H_0 : \theta = \theta_0$$

mot *mothypotesen*

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

eller mot en ensidig mothypotes $H_1 : \theta > \theta_0$ eller $H_1 : \theta < \theta_0$.

Om testet bekräftar mothypotesen säger vi att

- H_0 förkastas till förmån för H_1.
- θ är signifikant skild från θ_0.

I fallet då vi inte kan förkasta H_0 har vi *inte* visat att $\theta = \theta_0$.
Frågan om H_0 eller H_1 är sann lämnas öppen.

Konfidensintervallmetoden

Om vi bildar ett 95% konfidensintervall I_θ för θ kan vi direkt använda detta för att göra hypotestestet.

- Vi kan förkasta H_0 om intervallet inte innehåller θ_0.

Konfidensgraden vi använder när vi beräknar konfidensintervallet är $1-\alpha$, där α är signifikansnivån för hypotestestet.

Signifikansnivå

Felrisken eller signifikansnivån definieras som

$$\alpha = \mathsf{P}(H_0 \text{ förkastas } | H_0 \text{ sann}).$$

Om H_0 förkastas trots att H_0 är sann görs ett Typ 1 fel.

Om vi använder ett konfidensintervall är signifikansnivån $\alpha = \mathsf{P}(\theta \notin I_\theta | H_0) = \mathsf{P}(\theta_0 \notin I_\theta)$.

Teststorhet

Teststorhet

En teststorhet $T = T(X_1, \ldots, X_n)$ är en funktion av observationerna, och alltså en slumpvariabel. $T_{obs} = T(x_1, \ldots, x_n)$ är ett observerat värde av teststorheten för givna observationer.

I allmänhet innehåller teststorheten en skattning θ^* av parametern.

p-värde

p-värdet för testet definieras som sannolikheten under nollhypotesen att vi får ett värde T som är åtminstone lika "extremt" som det observerade värdet T_{obs}.

Förkasta H_0 om p-värdet är mindre än den valda signifikansnivån.

p-värden

Vi vill använda en storhet T som vi känner fördelningen av under H_0, så att vi kan beräkna p-värdet.

I fallet med normalfördelning med känd varians kan vi använda

$$T = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

har vi att T under H_0 är $\mathsf{N}(0,1)$-fördelad, så vi har att

$$p = \mathsf{P}(|T| \geq |T_{obs}|) = 2 \cdot \mathsf{P}(T \geq |T_{obs}|) = 2(1 - \Phi(|T_{obs}|)).$$

Vi förkastar H_0 om $p < \alpha$, och det exakta värdet på p är alltså den minsta signifikansnivå vi kan ha för att förkasta H_0.

Kritiskt område

Definition

Givet en signifikansnivå α definierar vi det kritiska området C_α som de värden på teststorheten T som leder till att man förkastar H_0 på nivån α.

Vi förkastar H_0 på nivån α om $p < \alpha$, vilket ekvivalent kan uttryckas som $|T| > T_{\alpha/2,H_0}$, där $T_{\alpha/2,H_0}$ är $\alpha/2$-kvantilen i T:s fördelning under H_0. Därför har vi $C_\alpha = \{T : |T| > T_{\alpha/2,H_0}\}$.

Exempel: Med test för väntevärdet hos normalfördelade data får vi:

- Om variansen är känd: T under H_0 är $\mathrm{N}(0,1)$-fördelad, förkasta H_0 på nivån α om $|T| > z_{\alpha/2}$.
- Om variansen är okänd: T under H_0 är $t(f)$-fördelad, förkasta H_0 på nivån α om $|T| > t_{\alpha/2}(f)$.

Styrka

Om vi inte förkasta H_0 trots att H_1 är sann gör vi ett Typ 2 fel.

Sannolikheten att *inte* göra ett typ 2 fel kallas för testets *styrka*, och ges av $1-\beta$, där $\beta = \mathsf{P}(\text{förkasta ej } H_0 | H_1 \text{ sann})$.

Typ 2 fel är lite svåra att hantera eftersom de beror på mothypotesen men ett sätt är att beräkna testets *styrkefunktion*

Definition

Vi vill testa $H_0 : \theta = \theta_0$ och har bestämt en teststorhet T och en signifikansnivå α. Testets styrkefunktion $\pi(\theta)$ definieras som

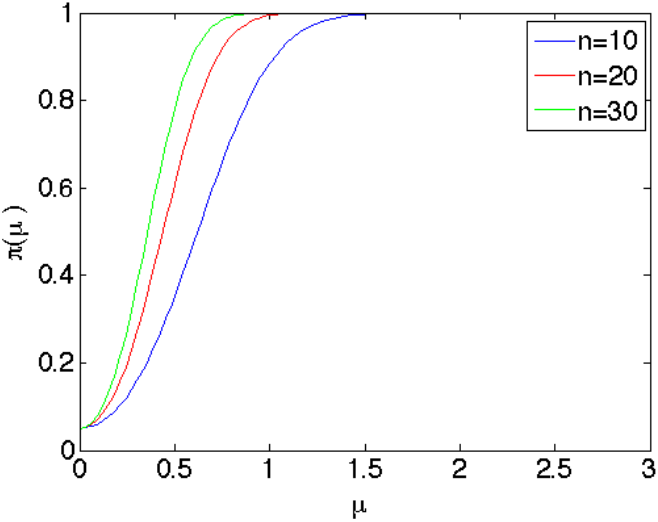
$$\pi(\theta) = \mathsf{P}(T \in C_\alpha | \theta) = \mathsf{P}(H_0 \text{ förkastas} | \text{ parametervärdet är } \theta)$$

Testets styrkefunktion ger alltså sannolikheten för att testet ger ett signifikant resultat om det sanna värdet är θ, vilket alltså är testets styrka vid värdet θ.

Sats

Antag att vi har n observationer av en normalfördelning $\mathsf{N}(\mu, \sigma^2)$ med känd varians och att vi vill testa $H_0 : \mu = \mu_0$. Testets styrka ges då av

- Om $H_1 : \mu > \mu_0$:
$$\pi(\mu) = 1 - \Phi\left(z_\alpha - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$$
- Om $H_1 : \mu < \mu_0$:
$$\pi(\mu) = \Phi\left(-z_\alpha - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$$
- Om $H_1 : \mu \neq \mu_0$:
$$\pi(\mu) = 1 + \Phi\left(-z_{\alpha/2} - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right) - \Phi\left(z_{\alpha/2} - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$$



Styrkefunktion för test av $H_0 : \mu = 0$ mot $H_1 : \mu \neq 0$ för N$(\mu, 1)$-fördelningen på nivå $\alpha = 0.05$.

Hur många observationer krävs?

Det vi kan se från styrkefunktionerna i de olika fallen är att testet har lättare att upptäcka en avvikelse $\mu - \mu_0$ om

- Risken för typ 1 fel, α, är stor.
- Antalet observationer, n, är stort.
- Variansen σ^2 är liten.

Om vi bestämmer oss för en viss avvikelse $\mu - \mu_0$ vi vill att testet ska upptäcka samt fixerar signifikansnivån α kan vi använda styrkefunktionen för att beräkna hur många observationer vi bör göra för att få en given styrka $1 - \beta$ på testet.

Sats

Antag att vi har normalfördelade observationer $\mathrm{N}(\mu, \sigma^2)$ och önskar testa $H_0 : \mu = \mu_0$. Om vi vill att testet ska ge utslag för avvikelsen $\mu - \mu_0$ med sannolikhet $1 - \beta$ ska antal observationer väljas som:

- Känd varians:

$$\text{Ensidigt test: } n = \frac{\sigma^2}{(\mu - \mu_0)^2}(z_\alpha + z_\beta)^2$$

$$\text{Tvåsidigt test: } n = \frac{\sigma^2}{(\mu - \mu_0)^2}(z_{\alpha/2} + z_\beta)^2$$

- Om variansen skattas med s^2:

$$\text{Ensidigt test: } n \approx \frac{\sigma^2}{(\mu - \mu_0)^2}(z_\alpha + z_\beta)^2 + z_\alpha^2/2$$

$$\text{Tvåsidigt test: } n \approx \frac{\sigma^2}{(\mu - \mu_0)^2}(z_{\alpha/2} + z_\beta)^2 + z_{\alpha/2}^2/2$$

Jämförelser

En vanlig situation är att man vill göra jämförelser av olika stickprov. Exempel på när detta kan vara av intresse är om

- Vi vill jämföra luftkvalitén i två städer.
- Vi vill undersöka effekten av ett nytt läkemedel.

Vi vill ofta se hur en *förklarande variabel* (tex temperatur) påverkar en *responsvariabel* (utbytet vid den kemiska reaktionen)

Vi kommer idag studera två typfall av jämförelser

- Oberoende stickprov (mätningar på två olika populationer)
- Stickprov i par (två mätningar på samma population)

Stickprov i par

En vanlig situation är att mätningarna uppkommer i par, till exempel om man vill studera

- Hur mycket rökare går upp i vikt när de slutar röka. Man mäter då vikten före och efter för varje person före och efter den slutar röka och jämförelsen sker för varje person.
- Systematiska skillnader mellan två mätmetoder och använder varje metod på var och ett av ett antal prover och jämför de två metoderna för varje prov.

Modellen vi ansätter är att vi har n observationer i två stickprov

$$X_1, X_2, \ldots, X_n \qquad\qquad Y_1, Y_2, \ldots, Y_n$$

För varje mätning bildar vi differensen som antas vara normalfördelad:

$$D_i = X_i - Y_i \sim \mathsf{N}(\Delta, \sigma^2)$$

Vi vill nu testa om $\Delta = 0$, vilket görs som vanligt för normalfördelade mätningar.

Oberoende stickprov

Vi antar att vi har två oberoende stickprov

- n_1 observationer $X_{11}, X_{12}, \ldots, X_{1n_1}$ från $\mathsf{N}(\mu_1, \sigma_1^2)$.
- n_2 observationer $X_{21}, X_{22}, \ldots, X_{2n_2}$ från $\mathsf{N}(\mu_2, \sigma_2^2)$.

Vi vill nu testa om vi kan anta att $\mu_1 = \mu_2$.

Inför $\theta = \mu_1 - \mu_2$ som vi skattar med $\theta^* = \bar{X}_1 - \bar{X}_2$. Testa

$$H_0 : \theta = 0,$$
$$H_1 : \theta \neq 0$$

Alternativt en ensidig mothypotes, $\theta > 0$ eller $\theta < 0$.

Vi skiljer på tre fall:

1. σ_1 och σ_2 är kända.
2. $\sigma_1 = \sigma_2 = \sigma$ där σ är okänd.
3. σ_2 och σ_2 är okända och ej säkert lika.

Jämförelse av variansen för oberoende stickprov

Om det inte är känt från försöksuppställningen kan vi behöva testa om vi kan anta $\sigma_1 = \sigma_2$.

Vi undersöker detta genom att utföra hypotestestet

$$H_0 : \sigma_1 = \sigma_2$$
$$H_1 : \sigma_1 \neq \sigma_2$$

Definition

Om $V_1 \sim \chi^2(f_1)$ och $V_2 \sim \chi^2(f_2)$ är oberoende så gäller att

$$\frac{V_1/f_1}{V_2/f_2}$$

är $F(f_1, f_2)$-fördelad ("F-fördelad med f_1 och f_2 frihetsgrader")

Täthetsfunktionen för $F(f_1, f_2)$-fördelningen