Punktskattningar

Observationer $x_1, \ldots, x_n$ av oberoende stokastiska variabler $X_1, \ldots, X_n$ i vars fördelning finns parameter θ.

En punktskattning av θ,

$$\theta^* = \theta^*(x_1, \ldots, x_n),$$

modelleras av skattningsvariabeln (stickprovsvariabeln)

$$\theta^* = \theta^*(X_1, \ldots, X_n).$$

Exempel:

	Skattning	**Skattningsvariabel**	**Parameter**
Stickprovsmedelvärde:	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$m = \mathsf{E}[X]$
Stickprovsvarians:	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$	$\sigma^2 = \mathrm{V}(X)$
Stickprovsstandardavvikelse:	$s = \sqrt{s^2}$	$S = \sqrt{S^2}$	$\sigma = \mathrm{D}(X)$

Egenskaper för skattningar:

Väntevärdesriktighet: Skattningen $\theta^*(x_1, \ldots, x_n)$ är en väntevärdesriktig skattning av θ om $\mathsf{E}[\theta^*(X_1, \ldots, X_n)] = \theta$.

Effektivitet: $\mathrm{V}(\theta^*(X_1, \ldots, X_n))$ liten

Konsistens (asymptotisk precision): $\mathsf{P}(|\theta^*(X) - \theta| > \epsilon) \to 0$ då $n \to \infty$.

Notera att begreppet effektivitet är meningsfullt för jämförelser av olika skattningsmetoder av samma parameter θ. Mindre varians är bättre!

Exempel: Opinionsundersökning (oändlig population). Skattningen $p^*(x) = x/n$ modelleras med $p^*(X) = X/n$ där X är $\mathrm{Bin}(n, p)$. Då är

$$\mathsf{E}[p^*(X)] = \mathsf{E}[X/n] = \frac{1}{n}\mathsf{E}[X] = \frac{1}{n}np = p,$$

dvs. skattningen är väntevärdesriktig, och

$$\mathrm{V}(p^*(X)) = \mathrm{V}(X/n) = \frac{1}{n^2}\mathrm{V}(X) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n} \to 0$$

då $n \to \infty$. Med Tjebychevs olikhet får vi

$$\mathsf{P}(|p^*(X) - p| > \epsilon) \leq \frac{\mathrm{V}(p^*(X))}{\epsilon^2} = \frac{p(1-p)}{\epsilon^2 n} \to 0$$

då n$\to \infty$ att skattningen är konsistent.

Med siffror, $n = 2854$ interjuer och $x = 1085$ sympatisörer, är skattningen

$$p^*(x) = \frac{x}{n} = \frac{1085}{2854} = 0.380.$$

Standardavvikelsen för $p^*(X)$ är

$$\mathrm{D}\left(p^*(X)\right) = \sqrt{\mathrm{V}\left(p^*(X)\right)} = \sqrt{\frac{p(1-p)}{n}}$$

vilken skattas med

$$\sqrt{\frac{p^*(x)(1-p^*(x))}{n}} = \sqrt{\frac{0.380(1-0.380)}{2854}} = 0.00909.$$

Definition: En skattning av $\mathrm{D}\left(\theta^*(X_1,\ldots,X_n)\right)$ kallas *medelfelet* till $\theta^*(x_1,\ldots,x_n)$.

Exempel: Skattning av väntevärden och varianser Låt $x_1,\ldots,x_n$ vara ett stickprov på X med väntevärde $m = \mathsf{E}\left[X_i\right]$ och varians $\mathrm{V}\left(X\right) = \sigma^2$. Skattningen

$$m^*(x_1,\ldots,x_n) = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

som beskrivs av

$$m^*(X_1,\ldots,X_n) = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

är väntevärdesriktig ty

$$\mathsf{E}\left[m^*(X_1,\ldots,X_n)\right] = \mathsf{E}\left[\overline{X}\right] = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\underbrace{\mathsf{E}\left[X_i\right]}_{=m} = m.$$

Vidare så är

$$\mathrm{V}\left(m^*(X_1,\ldots,X_n)\right) = \mathrm{V}\left(\overline{X}\right) = \mathrm{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \{\text{oberoende}\} = \frac{1}{n^2}\sum_{i=1}^{n}\underbrace{\mathrm{V}\left(X_i\right)}_{=\sigma^2} = \frac{\sigma^2}{n}$$

vilket med Tjebychevs olikhet ger att skattningen är konsistent.

Skattningen av σ^2, s^2, beskrivs av S^2 som ger väntevärdesriktiga skattningar: $\mathsf{E}\left[S^2\right] = \sigma^2$. Det är en nyttig övning att visa detta med räknelagarna!

Generella skattningsmetoder

Hur väljer man sina skattningar?

- Maximum likelihood-metoden (maximi-metoden, ML-metoden)
- Minsta kvadrat-metoden (MK-metoden)

I fortsättningen, låt $x_1,\ldots,x_n$ vara ett stickprov. Antag att i fördelningen för $(X_1,\ldots,X_n)$ ingår parametrarna $\theta_1,\ldots,\theta_r$.

Maximum likelihood-metoden Skattningarna $\theta_1^*,\ldots,\theta_r^*$ är de värden på $\theta_1,\ldots,\theta_r$ som maximerar *trolighetsfunktionen*

$$L(\theta_1,\ldots,\theta_r) = f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$$

vid kontinuerliga fördelningar, eller

$$L(\theta_1,\ldots,\theta_r) = \mathsf{P}\left(X_1 = x_1,\ldots,X_n = x_n\right)$$

vid diskreta.

Filosofin bakom Maximum likelihood-metoden är att vårt utfall (stickprovet) bestäms av parametrarna i de ingående fördelningarna. Skattningen av dessa parametrar är de värden som gör det observerade så troligt som möjligt.

Tips! Ofta är det enklare att maximera $\log L(\theta_1, \ldots, \theta_r)$.

Exempel: Binomialfördelningen. Låt X är $\mathrm{Bin}(n, p)$ med observerat utfall x.

$$L(p) = \mathsf{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Logaritmering ger

$$\ln L(p) = \ln \left(\binom{n}{x} \right) + x \ln p + (n-x) \ln(1-p).$$

Derivering bestämmer maximum:

$$\frac{d}{dp} \ln L(p) = 0 + x\frac{1}{p} - (n-x)\frac{1}{1-p} = 0$$

ger skattningen $p = x/n$. Detta gäller för alla utfall x och $p^*(x) = x/n$ och vi skriver skattningsvariablen som $p^*(X) = X/n$.

(Skall detta göras ordentligt skall man bestämma 2:a-derivatan för att se att det är ett maximum man fått fram, samt även undersöka randpunkterna.)

Minsta kvadrat-metoden Kräver inga fördelningsantaganden på $X_1, \ldots, X_n$, bara att vi kan modellera deras väntevärden. Skattningarna $\theta_1^*, \ldots, \theta_r^*$ de värden på $\theta_1, \ldots, \theta_r$ som minimerar *minsta kvadrat-avstånden* mellan observationer och väntevärden:

$$Q(\theta_1, \ldots, \theta_r) = \sum_{i=1}^{n} (x_i - \mathsf{E}[X_i])^2.$$

Även här minimeras Q oftast med hjälp av derivering.

Exempel: Opinionsundersökning. $\mathsf{E}[X] = np$. Minimering av $Q(p) = (x - np)^2$ ger $p = x/n$. MK-skattningen av p är således $p^*(x) = x/n$.