

Sannolikhet och statistik XIV

Johan Jonasson

May 2019

Maximum likelihood (ML): Punktskatta θ genom att maximera sannolikheten att få de data man fick.

Mer precist: Om X_1, \dots, X_n är ett sp på en sv med en frekvn $p(x) = p_\theta(x)$ som beror av θ och vi obs $X_1 = x_1, \dots, X_n = x_n$, finn det θ som max

$$\prod_{k=1}^n p_\theta(x_k) =: L(\theta; x_1, \dots, x_n).$$

Detta maximum, skrivet $\hat{\theta}$, kallas för ML-skattningen av θ .

Exempel: Låt $X \sim \text{Bin}(n, \theta)$. Finn ML-skattningen av θ .

$$L(\theta; x) = p_\theta(x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}.$$

Om $x = 0$ är L avtagande, så max fås då $\theta = 0$. Om $x = n$ är L växande och fås max då $\theta = 1$. Annars, lös

$$L'(\theta; x) \propto x\theta^{x-1}(1-\theta)^{n-x} - (n-x)\theta^x(1-\theta)^{n-1-x} = 0.$$

Detta ger

$$x(1 - \theta) - (n - x)\theta = 0.$$

Lösningen är $\theta = x/n$. Sammantaget

$$\hat{\theta} = \frac{x}{n}.$$

□

Obs som underlättar: $L(\theta; x_1, \dots, x_n)$ maximeras precis då $\ell(\theta; x_1, \dots, x_n) := \ln L(\theta; x_1, \dots, x_n)$ maximeras.

Exempel: Låt $X \sim \text{Geo}(\theta)$. ML-skatta θ . Vi har $L(\theta; x) = \theta(1 - \theta)^{x-1}$, så

$$\ell(\theta; x) = \ln \theta + (x - 1) \ln(1 - \theta).$$

Vi får

$$\ell'(\theta; x) = \frac{1}{\theta} - \frac{x-1}{1-\theta} = 0$$

vilket ger $\theta = 1/x$. ML-skattningen av θ är alltså

$$\hat{\theta} = \frac{1}{x}.$$

Exempel: Låt X_1, \dots, X_n vara ett sp på $Poi(\theta)$. Vi får $L(\theta) = \prod_k e^{-\theta} \theta^{x_k} / x_k!$. Alltså

$$\ell(\theta) = -n\theta + \left(\sum_k x_k \right) \ln \theta - \sum_k \ln(x_k!).$$

Alltså

$$\ell'(\theta) = -n + \frac{\sum_k x_k}{\theta} = 0.$$

Detta ger

$$\hat{\theta} = \frac{1}{n} \sum_k x_k = \bar{x}.$$

□

För kontinuerliga fördelningar maximerar vi istället tätheten av stickprovet som fkn av θ .

Exempel: Låt X_1, \dots, X_n vara ett sp på $N(\mu, \sigma^2)$. ML-skatta (μ, σ^2) .

$$L(\mu, \sigma) = \prod_k \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x_k - \mu)^2}{\sigma^2}} = \frac{1}{\sigma^n (2\pi)^{n/2}} e^{-\frac{1}{2} \sum_k \frac{(x_k - \mu)^2}{\sigma^2}}.$$

Vi får

$$\ell(\mu, \sigma) = -\frac{n}{2} \ln(2\pi) - n \ln \sigma - \frac{1}{2} \frac{\sum_k (x_k - \mu)^2}{\sigma^2}.$$

Alltså

$$\frac{\partial \ell}{\partial \mu} = \frac{\sum_k (x_k - \mu)}{\sigma^2}$$

som är 0 då $\mu = \bar{x}$. Alltså

$$\hat{\mu} = \bar{x}.$$

Vidare är

$$\frac{\partial \ell}{\partial \sigma}(\hat{\mu}, \sigma) = -\frac{n}{\sigma} + \frac{\sum_k (x_k - \hat{\mu})^2}{\sigma^3}.$$

Detta är 0 då $\sigma^2 = (1/n) \sum_k (x_k - \hat{\mu})^2$, dvs

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2.$$

□

Hypotesprövning/tester: Är slanten rättvis? Den neutrala hypotesen är att $p = 1/2$, men vi misstänker att $p \neq 1/2$. Vi vill samla data för att avgöra.

Filosofi: Utgå från att den neutrala hypotesen är sann. Om data utifrån detta antagande får värden som är osannolika och samtidigt skulle varit mer sannolika om den alternativa hypotesen är sann, drar vi slutsatsen att den senare är den sanna. (All analys bör göras innan data samlas in.) (Ingen exakt vetenskap, men oftast naturligt val.)

I vårt slantexempel: Vi vill testa *nollhypotesen*

$$H_0 : p = \frac{1}{2}$$

mot *alternativhypotesen*

$$H_A : p \neq \frac{1}{2}.$$

Fixera ett litet tal $\alpha > 0$ (ofta $\alpha = 0.05$ eller $\alpha = 0.01$). Beteckna data med \mathcal{X} och bestäm en mängd B_0 s.a.:

$$\mathbb{P}_{H_0}(\mathcal{X} \notin B_0) = \alpha$$

och $\mathcal{X} \notin B_0$ talar till förmån för H_A över H_0 . Om det sedan visar sig att $\mathcal{X} \notin B_0$ förkastar vi H_0 till förmån för H_A på signifikansnivån α . Om $\mathcal{X} \in B_0$ accepterar vi H_0 .

I exemplet: Data n slantsinglingar. Låt X vara antal klave. Om X avviker mycket från $n/2$ talar detta till förmån för H_A över H_0 . Enl CGS gäller under H_0 (dvs om H_0 är sann)

$$\frac{X - n/2}{\sqrt{n/4}} \approx N(0, 1).$$

Med $z_{\alpha/2} = \Phi^{-1}(1 - \alpha/2)$, gäller alltså

$$\mathbb{P}_{H_0} \left(\frac{X - n/2}{\sqrt{n/4}} \notin \pm z_{\alpha/2} \right) \approx \alpha.$$

Dvs

$$\mathbb{P}_{H_0} \left(X \notin \frac{n}{2} \pm z_{\alpha/2} \sqrt{\frac{n}{4}} \right) \approx \alpha$$

och vi förkastar H_0 till förmån för H_A om $X \notin n/2 \pm z_{\alpha/2} \sqrt{n/4}$.

Ex.vis med $n = 100$ och $\alpha = 0.05$ blir $z_{0.025} \sqrt{n/4} = 1.96 \sqrt{25} \approx 10$ och vi förkastar om X avviker med minst 10 från 50.

Med $n = 10000$ och $\alpha = 0.01$ blir $z_{0.005} \sqrt{n/4} = 2.58 \sqrt{2500} \approx 129$ och vi förkastar om X avviker med minst 129 från 5000.

Exempel: Låt X_1, \dots, X_n vara ett sp på $N(\mu, \sigma^2)$. Testa $H_0 : \mu = \mu_0$ mot $H_A : \mu \neq \mu_0$.

Om H_0 är sann bör \bar{X} ligga nära μ_0 och stora avvikelser från det talar för H_A .

Under H_0 :

$$\frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}.$$

Alltså

$$\mathbb{P}_{H_0} \left(\bar{X} \in \mu_0 \pm F_{t_{n-1}}^{-1}(1 - \alpha/2) \frac{s}{\sqrt{n}} \right) = 1 - \alpha$$

och vi förkastar H_0 till förmån för H_A på signnivå α om $\bar{X} \notin \mu_0 \pm F_{t_{n-1}}^{-1}(1 - \alpha/2) s/\sqrt{n}$.

Konkret fall: Marmeladburkar ska innehålla 400 g. Vi misstänker att det inte stämmer. Vi väljer 10 burkar på måfå. Rimligt att anta att burkars vikt är $N(\mu, \sigma^2)$ -fördelade.

Vi vill testa $H_0 : \mu = 400$ mot $H_A : \mu \neq 400$. Vi förkastar H_0 till förmån för H_A på signnivå 5% om

$$\bar{X} \notin 400 \pm F_{t_9}^{-1}(0.975) \frac{s}{\sqrt{10}}.$$

Vi slår upp att $F_{t_9}^{-1}(0.975) \approx 2.26$ vilket ger att vi förkastar om

$$|\bar{X} - 400| \geq 0.71s.$$

OK, men kanske mer intresserade av $H_0 : \mu \geq 400$ mot $H_A : \mu < 400$. Bara värden \bar{X} klart mindre än 400 som talar för H_A . Gör ensidigt test.

Generös mot H_0 : testa $H_0 : \mu = \mu_0$ mot $H_A : \mu < \mu_0$. Utnyttja att

$$\mathbb{P}_{H_0} \left(\bar{X} \leq \mu_0 - F_{t_{n-1}}^{-1}(1 - \alpha) \frac{s}{\sqrt{n}} \right) = \alpha$$

Förkasta H_0 på signnivå α om $\bar{X} < \mu_0 - F_{t_{n-1}}^{-1}(1 - \alpha)s/\sqrt{n}$.

För marmeladburkarna förkastar vi $H_0 : \mu = 400$ mot $H_A : \mu < 400$ på signnivå 5% om

$$\bar{X} < 400 - F_{t_9}^{-1}(0.95) \frac{s}{\sqrt{10}} \approx 400 - 0.58s.$$