

13 Bayesian reliability analysis

Huvudsyftet med den Bayesianska ansatsen till statistisk inferens är att på ett objektivt sätt inkorporera subjektiv information (t.ex en tillförlitlighetsteknikers erfarenhet av en liknande komponent som den vi studerar). I det som följer ger jag en väldigt pragmatisk egen syn på det Bayesianska tillvägagångssättet. Jag följer alltså inte alls lärobokens (och förmodligen inte heller någon annan boks) framställning av Bayesiansk metodik.

Tänk dig att du har provat n enheter och mätt upp livstiderna $t_1, \dots, t_n$. Antag för enkelhets skull att modellen är $T \sim \exp(\lambda)$. Trolighetsfunktionen är då

$$l(\lambda) = \lambda^n e^{-\lambda \sum_i t_i}$$

och dess maximum uppnås för

$$\hat{\lambda} = \frac{n}{\sum_i t_i} = \frac{1}{\bar{t}}$$

Detta är vår trolighetsskattning om vi ej tar med den förhandsinformation om enheten som vi tänker att vi har.

Antag att vi på förhand vet att $\lambda \approx \lambda_0$. För att kunna inkorporera denna information i trolighetsfunktionen, måste vi värdera dess säkerhet. I bästa fall har vi redan ett antal mätningar på denna komponent. Då kom vi kanske fram till skattningen $\hat{\lambda} = \lambda_0$ via trolighetsfunktionen

$$l_0(\lambda) = \lambda^{n_0} e^{-\lambda \sum_i s_i}$$

som alltså maximeras av $\lambda_0 = n_0 / \sum_i s_i$ och där vi tänker oss att vi gjort n_0 livslängdsmätningar och erhållit mätvärdena $s_1, \dots, s_{n_0}$. I en analys baserad på alla $n_0 + n$ mätningarna bildar man självklart trolighetsfunktionen

$$l_0(\lambda)l(\lambda) = \lambda^{n_0+n} e^{-\lambda(\sum_i s_i + \sum_i t_i)}$$

som har maximum i

$$\hat{\lambda} = \frac{n_0 + n}{\sum_i s_i + \sum_i t_i}$$

d.v.s ett genom medelvärde av alla mätningarna.

Men det typiska är nog att vi har tillgång till tidigare mätresultat för en liknande komponent. Dessa kanske inte fullt ut kan överföras till den nya

komponenten för vi har ju gjort en del ändringar i konstruktionen. Men vi kan kanske på basis av de mätningarna gissa att $\lambda \approx \lambda_0$ och värdera säkerheten i gissningen med ett ekvivalent antal mätningar n_0 . Då kan vi bakvägen konstruera en trolighetsfunktion $l_0(\lambda)$, som uppfyller

$$\lambda_0 = \underset{\lambda}{\operatorname{argmax}} l_0(\lambda)$$

så här

$$l_0(\lambda) = \lambda^{n_0} e^{-\lambda n_0 / \lambda_0}$$

Detta ska vi kalla vår à priori-trolighet (à priori är latin och betyder på förhand).

Multiplikation av à priori-troligheten $l_0(\lambda)$ med modell-troligheten $l(\lambda)$ ger oss troligheten

$$l_0(\lambda)l(\lambda) = \lambda^{n_0} e^{-\lambda n_0 / \lambda_0} \lambda^n e^{-\lambda \sum_i t_i} = \lambda^{n_0+n} e^{-\lambda(n_0/\lambda_0 + \sum_i t_i)}$$

som maximeras utav

$$\hat{\lambda} = \frac{n_0 + n}{n_0/\lambda_0 + \sum_i t_i}$$

Detta är vår à posteriori-skattning av λ och produkten $l_0(\lambda)l(\lambda)$ vår à posteriori-trolighet. (À posteriori är latin och betyder i efterhand, alltså då vi sett data.)

Detta sätt att inkorporera subjektiv information i den frekventistiska analysen kan man motivera m.h.a det Bayesianiska synsättet på statistik. I sådana sammanhang brukar man tala om à priori-tätheten för λ som vi ska beteckna $p(\lambda)$. Obs att detta är en sannolikhetstäthet i λ . I föreliggande exempel låter jag

$$p(\lambda) \propto \lambda^{n_0} e^{-\lambda n_0 / \lambda_0}$$

där proportionalitetskonstanten $c(n_0, \lambda_0)$ ska väljas så att

$$1 = \int_0^\infty p(\lambda) d\lambda = c(n_0, \lambda_0) \int_0^\infty \lambda^{n_0} e^{-\lambda n_0 / \lambda_0} d\lambda$$

Obs att man i den Bayesianiska sinnevärlden gärna definierar andra à priori-tätheter. Det enda man egentligen kräver är att den ska spegla ens förhands-tro. Fast jag menar ju att det finns bra argument för att välja à priori-tätheten så att den passar modellen, så som jag gjort här. Obs också att

i den analys vi ska göra spelar det exakta värdet av konstanten $c(n_0, \lambda_0)$ ingen som helst roll. Vi behöver alltså inte räkna ut den.

Vi har också en modell för våra mätresultat. Den betecknas $p(t_1, \dots, t_n | \lambda)$ och är

$$p(t_1, \dots, t_n | \lambda) = \lambda^n e^{-\lambda \sum_i t_i}$$

(J.f.r trolighetsfunktionen som vi skrev upp ovan.)

M.h.a följande generalisering av Bayes formel

$$p(\lambda | t_1, \dots, t_n) \propto p(\lambda)p(t_1, \dots, t_n | \lambda)$$

($p(\lambda | t_1, \dots, t_n)$ kallas för λ :s à posteriori-täthet) räknar vi ut att

$$p(\lambda | t_1, \dots, t_n) \propto \lambda^{n_0+n} e^{-\lambda(n_0/\lambda_0 + \sum_i t_i)}$$

(Detta är ju inget annat än produkten $l_0(\lambda)l(\lambda)$ av de två troligheter vi definierade ovan.)

Medelst maximering av à posteriori-tätheten erhåller vi vad man kan kalla à posteriori-skattningen av λ . Att den är

$$\hat{\lambda} = \frac{n_0 + n}{n_0/\lambda_0 + \sum_i t_i}$$

har vi redan konstaterat. Jag vill också gärna kalla λ_0 för à priori-skattningen av λ , eftersom

$$\operatorname{argmax}_{\lambda} l_0(\lambda) = \frac{n_0}{n_0/\lambda_0} = \lambda_0$$

À posteriori-konfidensintervall kan konstrueras. Motsvarigheten i den Bayesianska teorin kallas "credibility intervals". Man måste då beräkna à posteriori-tätheten för λ exakt. I det fall vi studerar är detta inte svårt, för det gäller att à posteriori-fördelningen för λ är $\text{Gamma}(\alpha, \beta)$ med

$$\alpha = n + n_0 + 1 \quad \text{och} \quad \frac{1}{\beta} = \frac{n_0}{\lambda_0} + \sum_i t_i$$

J.f.r kursens formelsamling (tänk på att R & H har en annorlunda parametrisering av Gammafördelningen). Låt $\lambda_{\alpha/2}$ och $\lambda_{1-\alpha/2}$ vara $\alpha/2$ - resp $(1 - \alpha/2)$ -kvantilerna m.a.p denna fördelning. Då gäller

$$\int_{\lambda_{\alpha/2}}^{\lambda_{1-\alpha/2}} p(\lambda | t_1, \dots, t_n) d\lambda = 1 - \alpha$$

och den Bayesianske statistikern kallar intervallet $[\lambda_{\alpha/2}, \lambda_{1-\alpha/2}]$ för ett kredibilitetsintervall med kredibilitet $(1 - \alpha)$.

Den frekventistiske statistikern kanske föredrar konfidensintervall framför kredibilitetsintervall. Han/hon använder då helt enkelt att

$$2\lambda \left(\frac{n_0}{\lambda_0} + \sum_i t_i \right) \sim \chi^2(2(n_0 + n))$$

Sammanfattningsvis kan man säga att i mitt pragmatiska sätt att ta in förhandsinformation i skattningarna, så konstruerar man en à priori-trolighet av samma typ som den trolighetsfunktion man har i den klassiska frekvensbaserade statistiken, multiplicerar ihop dessa och därefter medelst maximering söker trolighetsskattningen med förhandsinformation inkluderad. Det svåra steget är att korrekt värdera säkerheten i förhandsgissningen $\lambda \approx \lambda_0$ i termer av ett ekvivalent antal mätningar n_0 . Obs att n_0 faktiskt ej måste vara ett heltal. I konstruktionen av konfidensintervall får vi i så fall tänka på att χ^2 -fördelningen kan definieras i termer av Gammafördelningen och att den senare är väldefinierad även då n_0 ej är ett heltal.

Exempel 1 Antag att man i ett konkret fall tror att $\lambda \approx 2.4 \cdot 10^{-4}$ t.u⁻¹ (t.u är vår tidsenhet). Man har i $n = 10$ erhållit $\sum_i t_i = 53\,673$ t.u. Ett sätt att värdera säkerheten i à priori-gissningen är att helt enkelt relatera den till säkerheten i den vanliga trolighetsskattningen. Då kanske man kommer fram till att säkerheten i den gamla informationen svarar mot ca 15% av säkerheten totalt och då kan $n_0 \approx 2$ vara en vettig ansats, eftersom $2/(10+2) = 1/6 \approx 0.167 \approx 15\%$. Då följer att $n_0/\lambda_0 \approx 8\,333$ t.u, vilket ger à posteriori-skattningen

$$\hat{\lambda} = \frac{2 + 10}{8\,333 + 53\,673} = \frac{12}{62\,006} = 1.94 \cdot 10^{-4} \text{ t.u}^{-1}$$

J.f.r med

$$\hat{\lambda} = \frac{10}{53\,673} = 1.86 \cdot 10^{-4} \text{ t.u}^{-1}$$

vilket är vad vi hade fått om förhandstron ej hade tagits med. Vi ser att den, kanske, aningen pessimistiska förhandstron gör att skattningen av λ blir lite större än vad den skulle blivit utan medtagande av förhandsinformation.

Antag istället att man tror sig veta att $\lambda \leq 4.7 \cdot 10^{-4} \text{ t.u}^{-1}$ med relativt hög säkerhet. Vi jämför med formeln för ett uppåt begränsat konfidensintervall för λ :

$$\lambda \leq \frac{\chi_{2n_0, \alpha}^2}{2 \sum_i s_i}$$

Här är $\sum_i s_i = n_0/\lambda_0$. Således gäller

$$\chi_{2n_0, \alpha}^2 = 2 \frac{n_0}{\lambda_0} 4.7 \cdot 10^{-4} = 3.92 n_0$$

Obs att α är risken att $\lambda \leq 4.67 \cdot 10^{-4} \text{ t.u}^{-1}$ är falskt. Bedömer vi att $\alpha \approx 10\%$, får vi att $n_0 \approx 2$, ty $\chi_{4, 0.10}^2 = 7.78$.

Det frekventistiska 0.95-konfidensintervallet blir

$$\lambda \leq \frac{\chi_{20, 0.05}^2}{2 \cdot \sum_i t_i} = \frac{31.4104}{2 \cdot 53\,673} = 2.93 \cdot 10^{-4} \text{ t.u}^{-1}$$

om man ej tar hänsyn till förhandsinformation, och

$$\lambda \leq \frac{\chi_{24, 0.05}^2}{2 \cdot (n_0/\lambda_0 + \sum_i t_i)} = \frac{36.4150}{2 (8\,333 + 53\,673)} = 2.94 \cdot 10^{-4} \text{ t.u}^{-1}$$

om man tar hänsyn till den. I detta fallet, då förhandstron kanske var aningen pessimistisk, ser vi ingen nämnvärd skillnad i den övre gränsen i de två konfidensintervallen. Notera också att

$$\lambda_0 = 2.4 \cdot 10^{-4} < 2.93 \cdot 10^{-4}$$

och således tillhör det förstnämnda konfidensintervallet, så ingen kan heller hävda att förhandsgissningen är orimlig pessimistisk. Men får man ett resultat som starkt ifrågasätter förhandsgissningen, så ska den förmodligen omvärderas eller strykas helt. $\square$