Föreläsning 09

Statistisk inlärning och dataanalys (Kungliga Tekniska Högskolan)

Föreläsning 9

Punktskattning ger oss val för en eller flera parametrar i en modell med hjälp av datan i formen av ett stickprov. Vi vill inte dock alltid ha ett påstående om parametern. Istället skulle vi ofta svara en viss fråga om parametern. Exempelvis skulle vi kunna ställa frågan, "Givet ett stickprov $x_1, \ldots, x_n$ från en $f_X(x|\theta)$-population stämmer det att $\theta > 10$?" eller "Stämmer det att $\theta = 10$?" Eftersom vi vet inte värdet av θ och har istället bara data föredrar vi att svara frågan med ett svar tillsammans med en sannolikhet att svaret stämmer.

Generellt sett är en *hypotes* ett visst påstående om en populationparameter såsom $\mu = 0$, $\lambda > 3$ eller $\sigma^2 > 1$. Målet i hypotesprövning är att bestämma om hypotesen stöds av datan där datan ges i formen av ett stickprov. Låt $x_1, \ldots, x_n$ vara ett stickprov från en $f_X(x|\theta)$-population där θ är en parameter i parameterrummet Ω. Vi har en viss *nollhypotes* (eller kan man säga bara *noll* eller *grundhypotesen*) H_0 som påstår att θ tillhör en delmängd Ω_0 av Ω. Vi skriver nollhypotesen som

$$H_0 : \theta \in \Omega_0.$$

Som alternativ till nollhypotesen H_0 har vi en *mothypotes* (eller *alternativhypotes*) H_1 som består av alla möjliga värden av θ som inte tillhör nollhypotesen. Det vill säger att

$$H_1 : \theta \in \Omega_0^c,$$

där $\Omega_0^c := \Omega \setminus \Omega_0$ är komplementet av Ω_0 i parameterrummet. Vi skriver också $\Omega_1 = \Omega_0^c$.

I denna kurs ska vi säga att H_0 är *accepterad* om H_0 stöds av datan $x_1, \ldots, x_n$ och H_0 är *förkastad* om det är inte. Detta är lika att förkasta H_1 respektive acceptera H_1 eftersom H_0 och H_1 är komplementära mängder. Det är vanligtvis i statistik för man att bara varken förkasta H_0 eller inte förkasta H_0 eftersom vårt beslut är baserad på ett stickprov och därför det kan vara fel att säga "Ja, $\theta \in \Omega_0$." Med detta i åtanke föredrar många statistiker att bara säga "Nej, vi kan inte förkasta nollhypotesen när H_0 stöds av datan. Skillnaden mellan att acceptera och inte förkasta H_0 är inte oviktig men vi ska inte säga mer om det under denna kursen.

9.1 Frekventistisk hypotesprövning

Vi har sett redan i det föregående föreläsningen att man kan använda beslutsteori och Bayesiansk metoder för att ge ett svar till frågor om en parameter som "Givet ett stickprov $x_1, \ldots, x_n$ från en $f_X(x|\theta)$-population stämmer det att $\theta > 10$?" eller "Stämmer det att $\theta = 10$?" Det andra frågan är kanske inte så väldefinierad från det Bayesianska perspektivet. Det är eftersom Bayesianska statistiker beror inte på idéen att θ har ett riktigt värde men istället bara en fördelning. Från detta perspektiv kan man tror inte att det stämmer att även ställa frågan "Stämmer det att $\theta = 10$?"! Under det frekventistiska perspektivet tror man att θ har egentligen ett riktigt värdet som kan upptäckas. Därför är sådana frågor väldefinierade.

Exempel 9.1. Som ett först exempel betraktar vi leveranshastigheten av produkter från en viss fabrik. Fabriken har en baslinje leveranshastighet som de måste upprätthålla för att fortsätta vara i drift. Denna baslinjen är 3 leveranser med ett tusen artiklar per arbetsdag. Under COVID-19-pandeminlåsningarna levererade fabriken att baslinje hastigheten. Efter låsningarna var slut fabriken registrerade deras leveranshastighet för fem slumpmässiga dagar:

$$x_1 = 9, x_2 = 2, x_3 = 4, x_4 = 3, x_5 = 6.$$

Fabriken tror att deras genomsnittlig leveranshastighet har ökat efter pandemin. Under det frekventistiska perspektivet tror vi att detta väntevärde är en viss (men okänt) värde. Vi kan verifiera om påståendet stöds av datan. För att göra det modellerar vi leveranshastigheten med en $\text{Po}(\lambda)$-fördelning och tar nollhypotesen

$$H_0 : \lambda = 3.$$

Parameterrummet är bara $\Omega = \{\lambda \in \mathbb{R} : \lambda \geq 3\}$ eftersom 3 är fabrikens minst möjlig leveranshastighet och så alternativhypotesen är

$$H_1 : \lambda > 3.$$

Som ett först steg kan vi betrakta sannolikheten av datan om nollhypotesen stämmer. Det kan ge oss en idé om hur rimligt är nollhypotesen baserad på den observerade datan från perspektivet av sannolikhet. Eftersom vi skulle vilja betrakta sannolikheten av den genomsnittliga leveranshästigheten använder vi statistisken $\bar{x}$ och betrakta sannolikheten $P(\bar{x}|\lambda = 3)$. Därför behöver vi samplingfördelningen $f_{\bar{X}}(\bar{x}|\lambda)$ vilken kan beräknas att vara

$$f_{\bar{X}}(\bar{x}|\lambda) = \frac{(5\lambda)^{5\bar{x}} e^{-5\lambda}}{(5\bar{x})!} \qquad \text{för } \bar{x} \in \{0, \tfrac{1}{5}, \tfrac{2}{5}, \ldots\}.$$

För givet datan har vi att

$$f_{\bar{X}}(\bar{x}|\lambda = 3) = f_{\bar{X}}(4.8|\lambda = 3) = \frac{(15)^{5(4.8)} e^{-15}}{(5(4.8))!} \approx 0.0195.$$

Eftersom denna sannolikhet är liten är det mycket osannlikt att vi skulle ha sett datan om nollhypotesen stäms. Så kan vi förkasta nollhypotesen $\lambda = 3$. Likaså accepterar vi alternativhypotesen $\lambda > 3$ i stöd av fabrikens åsikt.

Med tänke på att fabriken kanske vill köra experimentet igen i framtiden skulle det vara mycket lättare för dem om vi kan ge en regel för att förkasta nollhypotesen så att det behövs inte att beräkna och betrakta en sådan sannolikhet varje gång.

Till exempel vet vi att $\mathrm{E}[\bar{X}|\lambda = 3] = 3$. Så möjligheten att se $\bar{x}$ mycket större än 3 är ganska liten. För att mindre risken att vi förkasta nollhypotesen när det stämmer är det en bra idé att bestämma att $\bar{x}$ är för stor när sannolikheten av $\bar{x}$ är liten. Därför en bra regel för att förkasta nollhypotesen är om sannolikheten givet $\lambda = 3$ att $\bar{x}$ är större än ett visst värde t är liten. Så skulle vilja vi hitta värdet t så att $P(\bar{X} \geq t|\lambda = 3)$ är liten. Vi får bestämma själv värdet av denna sannolikhet; dvs. vi får bestämma *signifikansnivån* av testet. Exempelvis, om fabriken skulle vilja ha signifikansnivån 0.05 behöver vi t så att $P(\bar{X} \geq t) = 0.05$. Vi beräknar värdet t med hjälp av formeln

$$\begin{aligned} P(\bar{X} \geq t|\lambda = 3) &= 1 - P(\bar{X} < t|\lambda = 3), \\ &= 1 - P(5\bar{X} < k|\lambda = 3), \qquad (\text{där } k \in \{0, 1, 2, \ldots\}) \\ &= 1 - \sum_{i=0}^{k-1} \frac{(15)^{5k} e^{-15}}{k!}, \end{aligned}$$

och se att för $t = 21/5 = 4.2$ är $P(\bar{X} \geq t) \approx 0.05$. Så har vi reglen

$$\text{Förkasta } H_0 : \lambda = 3 \text{ om } \bar{X} \geq 4.2.$$

Generellt sett är ett *hypotestest* (eller ett *test*) en regel som vi använda för att förkasta H_0 (och därför acceptera H_1). Mängden av stickprov $\boldsymbol{x}$ för vilka hypotesprövningen förkastar H_0 kallas *kritiskområdet* (eller *förkastningsområdet*). Kritiskområdet brukar betecknas som R. Exempelvis har vi kritiskområdet $R = \{\boldsymbol{x} \in \mathbb{R}^5 : \bar{x} \geq 4.2\}$ i Exempel 9.1. Hypotesprövningar formuleras ofta med en statistiska

$$W(\boldsymbol{X}) = W(X_1, \ldots, X_n)$$

som kallas för en *testvariabel*.

Givet att det finns många olika situationer där vi skulle vilja ha en hypotesprövning – exempelvis, när vi använder olika modeller — skulle det vara bra att ha en stor familj av hypotesprövning som fungerar för olika modeller och hypoteser. Nu betraktar vi en sådan familj.

9.2 Likelihood-kvottest

Ett sett att utvärdera en hypotes $H_0 : \theta \in \Omega_0$ är att jämföra likelihooden av hypotesen med likelihooden över hela parameterrummet. Om likelihooden att $\theta \in \Omega_0$ är tillräckligt liten relativ till likelihooden över hela parameterrummet har vi en bra anledning att förkasta H_0. Så vi skulle vilja betrakta provstatistiken

$$\lambda(\boldsymbol{x}) = \frac{\max_{\theta \in \Omega_0} L(\theta|\boldsymbol{x})}{\max_{\theta \in \Omega} L(\theta|\boldsymbol{x})}$$

som kallas för *likelihood-kvotstestvariabel* eller bara *likelihood-kvoten*.

Notisera att ML-skattningen $\hat{\theta}$ för datan $\boldsymbol{x}$ uppfyller

$$\hat{\theta} = \mathrm{argmax}_{\theta\in\Omega} L(\theta|\boldsymbol{x})$$

och därför är $\hat{\theta}$ ML-skattningen för θ om och endast om

$$\max_{\theta\in\Omega} L(\theta|\boldsymbol{x}) = L(\hat{\theta}|\boldsymbol{x}).$$

Så numeraren i likelihood-kvoten är bara likelihoodfunktionen evalueras på ML-skattningen. På samma sett kan vi beräkna täljaren som $L(\hat{\theta}_0|\boldsymbol{x})$ där $\hat{\theta}_0$ är ML-skattningen för den begränsade likelihood optimeringsproblem

$$\hat{\theta}_0 = \mathrm{argmax}_{\theta\in\Omega_0} L(\theta|\boldsymbol{x}).$$

Eftersom vi vet redan hur att beräkna ML-skattningar för många olika modeller kan vi beräkna även sådana likelihoodkvoten!

Medan det händer ofta att Ω_0 är en sluten mängd (till exempel, $H_0: \mu = 3$ eller $H_0: \lambda \geq 5.7$) det kan hända att Ω_0 är öppen (exempelvis, $H_0: \lambda > 5.7$). För att ge ett likelihood-kvottest i sådana fall ersätter vi max med sup i likelihoodkvotet:

$$\lambda(\boldsymbol{x}) = \frac{\sup_{\theta\in\Omega_0} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Omega} L(\theta|\boldsymbol{x})}.$$

Notiser att täljaren av $\lambda(\boldsymbol{x})$ är alltid mindre än eller lika med nämnaren; dvs. $0 \leq \lambda(\boldsymbol{x}) \leq 1$ när $L(\hat{\theta}|\boldsymbol{x}) > 0$. När $\lambda(\boldsymbol{x})$ är liten följer det att vi kan hitta parametrar i $\Omega \setminus \Omega_0$ som ger en ganska mycket bättre likelihood än alla parametrarna i Ω_0. Så vi kan formulera den *likelihood-kvottestet*

$$\text{Förkasta } H_0: \theta \in \Omega_0 \text{ när } \lambda(\boldsymbol{x}) \leq c \text{ för någon } c \in (0,1).$$

Precis hur vi får att bestämma själv gränsen 0.05 i Exempel 9.1 får vi även att bestämma själv valet för $c \in (0,1)$ som vi använder i likelihood-kvottestet.

Exempel 9.2. Låt $x_1, \dots, x_n$ vara ett stickprov från en $\mathrm{N}(\mu, 1)$-population. Vi skulle vilja testa

$$H_0: \mu = \mu_0 \qquad \text{mot} \qquad H_1: \mu \neq \mu_0.$$

För detta test är $\Omega_0 = \{\mu_0\}$ och $\Omega_0^c = \mathbb{R} \setminus \{\mu_0\}$ och likelihoodkvotet

$$\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{\sup_{\mu\in\Omega_0} L(\mu|\boldsymbol{x})}{\sup_{\mu\in\Omega} L(\mu|\boldsymbol{x})}, \\
&= \frac{L(\mu_0|\boldsymbol{x})}{\sup_{\mu\in\mathbb{R}} L(\mu|\boldsymbol{x})}, \\
&= \frac{L(\mu_0|\boldsymbol{x})}{L(\hat{\mu}|\boldsymbol{x})}, \\
&= \frac{f_{\boldsymbol{X}}(\boldsymbol{x}|\mu_0)}{f_{\boldsymbol{X}}(\boldsymbol{x}|\hat{\mu})},
\end{aligned}$$

där $\hat{\mu}$ är ML-skattningen för μ givet datan. Vi såg i Exempel 5.2 att $\hat{\mu} = \bar{x}$. Så har vi att

$$\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{f_{\boldsymbol{X}}(\boldsymbol{x}|\mu_0)}{f_{\boldsymbol{X}}(\boldsymbol{x}|\hat{\mu})}, \\
&= \frac{\frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i - \mu_0)^2\right)}{\frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i - \bar{x})^2\right)}, \\
&= \exp\left(-\frac{1}{2}\left(\sum_{i=1}^n (x_i - \mu_0)^2 - \sum_{i=1}^n (x_i - \bar{x})^2\right)\right), \\
&= \exp\left(-\frac{1}{2} n(\bar{x} - \mu_0)^2\right),
\end{aligned}$$

där vi har använt formeln

$$\sum_{i=1}^n (x_i - \mu_0)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2.$$

En likelihood-kvottest kan då formuleras som

$$\text{Förkasta } H_0 : \theta = \theta_0 \text{ om } e^{-\frac{n}{2}(\bar{x}-\mu_0)^2} \leq c \text{ gör en given } c \in (0,1).$$

Med en liten algebra ser vi att testet är lika till

$$\text{Förkasta } H_0 : \mu = \mu_0 \text{ om } |\bar{x} - \mu_0| \geq c',$$

där $c' = \sqrt{-\frac{2}{n}\log(c)}$. Eftersom $c \in (0,1)$ är godtycklig kan vi ta vilken c' som helst i $(0,\infty)$. För detta test förkastar vi H_0 när stickprovsmedelvärdet är tillräcklig lång borta från μ_0. Testvariabeln är

$$W(\boldsymbol{x}) = |\bar{x} - \mu_0|,$$

där vi förkasta H_0 när $W(\boldsymbol{x})$ är stor. Kritiskt området är

$$R = \{\boldsymbol{x} : |\boldsymbol{x} - \mu_0| \geq c'\},$$

där $c' \in (0,\infty)$ är fixt för testet.

Exempel 9.3. Låt $X_1, \ldots, X_n$ vara oberoende och Exp(1)-läge familj med läge parameter θ; dvs.

$$f_X(x|\theta) = f_Z(x-\theta) \qquad \text{där } Z \sim \text{Exp}(1).$$

Så har vi att

$$\begin{aligned} f_X(x|\theta) &= \begin{cases} e^{-(x-\theta)} & x \geq \theta, \\ 0 & x < \theta, \end{cases} \\ &= e^{-(x-\theta)}\mathbf{1}_{\{y:y\geq\theta\}}(x). \end{aligned}$$

Vi skulle vilja testa

$$H_0 : \theta \leq \theta_0 \qquad \text{mot} \qquad H_1 : \theta > \theta_0,$$

för någon θ_0. Parameterrummet är $\Omega = \mathbb{R}$ och parametrarna under H_0 är $\Omega_0 = \{\theta \in \mathbb{R} : \theta \leq \theta_0\} = (-\infty, \theta_0]$. Likelihoodfunktionen för θ är

$$\begin{aligned} L(\theta|\boldsymbol{x}) &= \prod_{i=1}^{n} f_X(x_i|\theta), \\ &= \prod_{i=1}^{n} e^{-(x-\theta)}\mathbf{1}_{\{y:y\geq\theta\}}(x_i), \\ &= e^{-\sum_{i=1}^{n}(x_i-\theta)}\mathbf{1}_{\{y:y\geq\theta\}}(\min_i(x_i)), \\ &= e^{n\theta - \sum_{i=1}^{n} x_i}\mathbf{1}_{\{\eta:\eta\leq\min_i(x_i)\}}(\theta), \end{aligned}$$

eftersom $\min_i(x_i) \geq \theta$ om och endast om $\theta \leq \min_i(x_i)$. Det följer att $L(\theta|\boldsymbol{x})$ är lika med 0 för $\theta > \min_i(x_i)$ och strikt växande på $(-\infty, \min_i(x_i)]$. Så ML-skattningen $\hat{\theta}$ för θ (över Ω) är lika med $\min_i(x_i)$. Då har vi att

$$\sup_{\theta\in\Omega} L(\theta|\boldsymbol{x}) = L(\min_i(x_i)|\boldsymbol{x}) = e^{-\sum_{i=1}^{n} x_i + n\min_i(x_i)}.$$

Under H_0 beror ML-skattningen $\hat{\theta}_0$ på om $\min_i(x_i)$ är mindre eller större än θ_0. Om $\min_i(x_i) \geq \theta_0$ då har vi att

$$\sup_{\theta\in\Omega_0} L(\theta|\boldsymbol{x}) = e^{-\sum_{i=1}^{n} x_i + n\theta_0}.$$

Om $\min_i(x_i) < \theta_0$ har vi att $\hat{\theta} = \min_i(x_i) \in \Omega_0$ och därför

$$\sup_{\theta\in\Omega_0} L(\theta|\boldsymbol{x}) = \sup_{\theta\in\Omega} L(\theta|\boldsymbol{x}).$$

Det följer att

$$\begin{aligned} \lambda(\boldsymbol{x}) &= \begin{cases} 1 & \min_i(x_i) < \theta_0, \\ \frac{e^{-\sum_{i=1}^{n} x_i + n\theta_0}}{e^{-\sum_{i=1}^{n} x_i + n\min_i(x_i)}} & \min_i(x_i) \geq \theta_0, \end{cases} \\ &= \begin{cases} 1 & \min_i(x_i) < \theta_0, \\ e^{-n(\min_i(x_i)-\theta_0)} & \min_i(x_i) \geq \theta_0, \end{cases} \end{aligned}$$

Så vi har likelihood-kvottestet

$$\text{Förkasta } H_0 : \theta \leq \theta_0 \text{ om } e^{-n(\min_i(x_i)-\theta_0)} \leq c \text{ för någon } c \in (0,1).$$

Notisera att vi kan ignorera fallet $\min_i(x_i) < \theta_0$ eftersom $c < 1$. Pa samma sett som i det föregående exemplet kan vi formulerar om hypotestestet för att få

$$\text{Förkasta } H_0 : \theta \leq \theta_0 \text{ om } \min_i(x_i) - \theta_0 \geq c' \text{ för någon } c' > 0.$$

Här har vi $c' = -\frac{1}{n}\log(c)$. Kritiskt området för testet är då $R = \{\boldsymbol{x} : \min_i(x_i) - \theta_0 \geq c'\}$.

Vi kan tillämpa testet till en situation som man kan modellera med en exponentialfördelning. Exponentialfördelningar användas ofta att modellera tiden mellan evenemang som är poissonfördelad. En poissonmodell kan vara en enkel modell för vulkanutbrottning. Eftersom det kan ta lite tid för en vulkan att bli kritiskt aktiv igen efter varje vulkanutbrottning stämmer det att modellera tiden mellan vulkanutbrottningarna med en skiftat exponentiell modell. Åren mellan vulkanutbrotten av vulkanen Mt. Sankt Helens i Washington, USA registrerades:

$$15, 9, 59, 18, 5, 41, 3, 7, 5, 7, 4.$$

Om vi modellerar tiden mellan vulkanutbrotten som $X = Z + \theta$ där $Z \sim \text{Exp}(1)$ har vi att $\min_i(x_i) = 3$. Om vi tar ett ganska litet värde för c, exempelvis, $c = 0.05$, har vi att $c' \approx 0.27$. Så testet ska förkasta $H_0 : \theta \leq \theta_0$ när $\theta_0 = 2.5$ men acceptera $H_0 : \theta \leq \theta_0$ när $\theta = 2.75$.

I båda exemplar testvariablen berrode bara på en tillräcklig statistika för parametern ($\bar{x}$ för μ och $\min_i(x_i)$ för θ). Denna fenomen gäller mer generellt.

Sats 9.1. *Om $T(\boldsymbol{X})$ är en tillräcklig statistika för θ och $\lambda^*(t)$ och $\lambda(\boldsymbol{x})$ är likelihood-kvottestvariabler baserade på T respektive $\boldsymbol{X}$ gäller*

$$\lambda^*(T(\boldsymbol{X})) = \lambda(\boldsymbol{X})$$

för alla $\boldsymbol{x}$.

Bevis. Låt $X_1, \ldots, X_n$ vara oberoende och likafördelade med täthetsfunktion (eller sannolikhetsfunktion) $f_X(x|\theta)$. Eftersom $T(\boldsymbol{X})$ är en tillräcklig statistika för θ säger Sats 3.6 (Faktorisering satsen) att det finns funktioner $h(\boldsymbol{x})$ och $g(t|\theta)$ så att

$$f_{\boldsymbol{X}}(\boldsymbol{x}|\theta) = h(\boldsymbol{x})g(T(\boldsymbol{x})|\theta).$$

Det följer att

$$\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{\sup_{\theta\in\Omega_0} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Omega} L(\theta|\boldsymbol{x})}, \\
&= \frac{\sup_{\theta\in\Omega_0} f_{\boldsymbol{X}}(\boldsymbol{x}|\theta)}{\sup_{\theta\in\Omega} f_{\boldsymbol{X}}(\boldsymbol{x}|\theta)}, \\
&= \frac{\sup_{\theta\in\Omega_0} h(\boldsymbol{x})g(T(\boldsymbol{x})|\theta)}{\sup_{\theta\in\Omega} h(\boldsymbol{x})g(T(\boldsymbol{x})|\theta)}, \\
&= \frac{\sup_{\theta\in\Omega_0} g(T(\boldsymbol{x})|\theta)}{\sup_{\theta\in\Omega} g(T(\boldsymbol{x})|\theta)}, \qquad (h(\boldsymbol{x}) \text{ är konstant och i både täljaren och numeraren}) \\
&= \frac{\sup_{\theta\in\Omega_0} L^*(\theta|T(\boldsymbol{x}))}{\sup_{\theta\in\Omega} L^*(\theta|T(\boldsymbol{x}))}, \qquad (g(T(\boldsymbol{x})|\theta) = f_{T(\boldsymbol{X})}(T(\boldsymbol{x})|\theta)) \\
&= \lambda^*(T(\boldsymbol{x})).
\end{aligned}$$

□

Med hjälp av Sats 9.1 kan beräkningarna i våra exemplar vara förenklas.

Exempel 9.4. Vi återgår till Exempel 9.2 där vi hade ett stickprov $x_1, \ldots, x_n$ från en $\text{N}(\mu, 1)$-population och ville testa

$$H_0 : \mu = \mu_0 \qquad \text{mot} \qquad H_1 : \mu \neq \mu_0.$$

Vi vet från Exempel 4.5 att $\bar{x}$ är en tillräcklig statistika. Vi vet även från Sektion 4.1.1 att $\bar{X} \sim \text{N}(\mu, \sigma^2/n)$. Det följer att

$$L(\mu|\bar{x}) = \frac{1}{(2\pi/n)^{1/2}} e^{-\frac{n}{2}(\mu-\bar{x})^2},$$

och så

$$\sup_{\mu\in\Omega} L(\mu|\bar{x}) = L(\bar{x}|\bar{x}) = \frac{1}{(2\pi/n)^{1/2}}e^{-\frac{n}{2}(\bar{x}-\bar{x})^2} = \frac{1}{(2\pi/n)^{1/2}}.$$

Vi har också att

$$\sup_{\mu\in\Omega_0} L(\mu|\bar{x}) = \sup_{\mu\in\{\mu_0\}} L(\mu|\bar{x}) = L(\mu_0|\bar{x}) = \frac{1}{(2\pi/n)^{1/2}}e^{-\frac{n}{2}(\mu_0-\bar{x})^2}.$$

Därför har vi att

$$\lambda(\bar{x}) = \frac{\frac{1}{(2\pi/n)^{1/2}}e^{-\frac{n}{2}(\mu_0-\bar{x})^2}}{\frac{1}{(2\pi/n)^{1/2}}} = e^{-\frac{n}{2}(\mu_0-\bar{x})^2} = \lambda(\boldsymbol{x}).$$

Detta är detsamma testvariabel som vi hade i Exempel 9.2.

Exempel 9.5. I Exempel 9.3 hade vi ett stickprov $x_1, \ldots, x_n$ från en $f_X(x|\theta)$-population där $f_X(x|\theta)$ tillhör till en läge familj med standardtäthetsfunktion $f_Z(z) = e^{-z}$ och θ var den läge parametern. Vi såg att vi kan skriver den simultana fördelningen $f_{\boldsymbol{X}}(\boldsymbol{x}|\theta)$ som

$$f_{\boldsymbol{X}}(\boldsymbol{x}|\theta) = e^{-\sum_{i=1}^n x_i}e^{n\theta)}\mathbf{1}_{\{y:y\geq\theta\}}(\min_i(x_i)).$$

Så följer det från faktorisering satsen (Sats 3.6) att $\min_i(x_i)$ är en tillräcklig statistika för θ. Eftersom $X_i = Z_i + \theta$ för oberoende och Exp(1)-fördelade $Z_1, \ldots, Z_n$ och

$$\min_i(X_i) = \min_i(Z_i + \theta) = \min_i(Z_i) + \theta$$

har vi att $f_{\min_i(X_i)}(\min_i(x_i)) = f_{\min_i(Z_i)}(\min_i(x_i) - \theta)$. Det är bra övning att visa att samplingfördelningen $f_{\min_i(Z_i)}(\min_i(z_i))$ är

$$f_{\min_i(Z_i)}(\min_i(z_i)) = \begin{cases} ne^{-n\min_i(z_i)} & 0 \leq \min_i(z_i), \\ 0 & 0 > \min_i(z_i). \end{cases}$$

Det följer att

$$f_{\min_i(X_i)}(\min_i(x_i)) = \begin{cases} ne^{-n(\min_i(x_i)-\theta)} & \theta \leq \min_i(x_i), \\ 0 & \theta > \min_i(x_i). \end{cases}$$

Eftersom $ne^{-n(\min_i(x_i)-\theta)}$ är stikt växande för $\theta \leq \min_i(x_i)$ följer det är ML-skattningen $\hat{\theta} = \min_i(x_i)$ och ML-skattningen $\hat{\theta}_0$ under H_0 är $\min_i(x_i)$ om $\min_i(x_i) < \theta_0$ och θ_0 om $\min_i(x_i) \geq \theta_0$. Så får vi att

$$\begin{aligned}\lambda(\min_i(x_i)) &= \begin{cases} 1 & \min_i(x_i) < \theta_0, \\ \frac{ne^{-n(\min_i(x_i)-\theta_0)}}{ne^{-n(\min_i(x_i)-\min_i(x_i))}} & \min_i(x_i) \geq \theta_0, \end{cases} \\ &= \begin{cases} 1 & \min_i(x_i) < \theta_0, \\ e^{-n(\min_i(x_i)-\theta_0)} & \min_i(x_i) \geq \theta_0, \end{cases} \\ &= \lambda(\boldsymbol{x})\end{aligned}$$

där $\lambda(\boldsymbol{x})$ är likelihood-kvoten som vi beräknas i Exempel 9.3.