EXAM ESTIMATION - DETECTION

26/10/2017

SOLUTIONS

Exercise 1: Estimation

1. Maximum likelihood estimation

   (a)

$$f(x_1,\cdots,x_n;\lambda)=\prod_{i=1}^n \lambda\exp(-\lambda x_i)=\lambda^n\exp\left(-\lambda\sum_{i=1}^n x_i\right)$$

$$\ln f(x_1,\cdots,x_n;\lambda)=n\ln(\lambda)-\lambda\sum_{i=1}^n x_i$$

$$\frac{\partial \ln f(x_1,\cdots,x_n;\lambda)}{\partial\lambda}=\frac{n}{\lambda}-\sum_{i=1}^n x_i\overset{!}{=}0\qquad\Longrightarrow\lambda=\frac{n}{\sum_{i=1}^n x_i}$$

$$\frac{\partial^2 \ln f(x_1,\cdots,x_n;\lambda)}{\partial\lambda^2}=-\frac{n}{\lambda^2}<0\qquad\Longrightarrow\quad\hat\lambda_{ML}=\frac{n}{\sum_{i=1}^n x_i}$$

   (b) Since $\varphi_X(t)=\mathbb{E}\left[\mathrm{e}^{itX}\right]=(1-it/\lambda)^{-1}$, it follows that

$$\varphi_{Y=\sum_{m=1}^n X_m}(t)=\mathbb{E}\left[\mathrm{e}^{it\sum_{m=1}^n X_m}\right]=\prod_{m=1}^n\mathbb{E}\left[\mathrm{e}^{itX_m}\right]=(1-it/\lambda)^{-n}$$

   and hence $T\sim\mathcal{G}(n,1/\lambda)$.

   (c) Since $\hat\lambda_{ML}=\frac{n}{Y}$ with $Y\sim\mathcal{G}(n,1/\lambda)$ and $\mathbb{E}[g(y)]=\int g(y)f_Y(y)dy$, we find

$$\mathbb{E}\left[\hat\lambda_{ML}\right]=n\int\frac{1}{y}\frac{1}{\Gamma(n)1/\lambda^n}y^{n-1}\exp\left(-\frac{y}{1/\lambda}\right)dy$$

$$=n\lambda\frac{\Gamma(n-1)}{\Gamma(n)}\underbrace{\int\frac{1}{\Gamma(n-1)1/\lambda^{n-1}}y^{n-2}\exp\left(-\frac{y}{1/\lambda}\right)dy}_{\int\mathcal{G}(n-1,1/\lambda)=1}=\frac{n}{n-1}\lambda$$

$$\Longrightarrow\text{biased:}\quad b(\hat\lambda_{ML})=\frac{1}{n-1}\lambda.$$

$$\mathbb{E}\left[\hat\lambda_{ML}^2\right]=n^2\int\frac{1}{y^2}\frac{1}{\Gamma(n)1/\lambda^n}y^{n-1}\exp\left(-\frac{y}{1/\lambda}\right)dy$$

$$=n^2\lambda^2\frac{\Gamma(n-2)}{\Gamma(n)}\underbrace{\int\frac{1}{\Gamma(n-2)1/\lambda^{n-2}}y^{n-3}\exp\left(-\frac{y}{1/\lambda}\right)dy}_{\int\mathcal{G}(n-2,1/\lambda)=1}=\frac{n^2}{(n-1)(n-2)}\lambda^2$$

$$\text{Var}\left[\hat\lambda_{ML}\right]=\frac{n^2}{(n-1)(n-2)}\lambda^2+\frac{n^2}{(n-1)^2}\lambda^2=\frac{n^2\lambda^2}{(n-1)^2(n-2)}\qquad\Longrightarrow\text{convergent}$$

   (CONTROL: the random variable $\hat\lambda_{ML}/n$ has an inverse Gamma distribution $\mathcal{IG}(n,\lambda)$ and hence mean and variance $\frac{n}{n-1}\lambda$ and $\frac{n^2\lambda^2}{(n-1)^2(n-2)}$, respectively.)

(d)

$$I(\lambda) = \mathbb{E}\left[-\frac{\partial^2 \ln f(x_1, \cdots, x_n; \lambda)}{\partial \lambda^2}\right] = \frac{n}{\lambda^2}$$

$$\text{bias: } \frac{db(\hat{\lambda}_{ML})}{d\lambda} = \frac{1}{n-1}$$

$$CRB(\lambda) = \frac{(1 - b'(\lambda))^2}{I(\lambda)} = \frac{n\lambda^2}{(n-1)^2}.$$

$\Longrightarrow$ $\hat{\lambda}_{ML}$ is not the efficient estimator for λ since 1) it is biased and 2) its variance exceeds the CRB

2. Estimator $\hat{\lambda}_2 = \frac{n-1}{\sum_{i=1}^n x_i}$

(a) It follows immediately from 1.(c) that

$$\mathbb{E}[\hat{\lambda}_2] = \lambda \quad \text{and} \quad \text{Var}[\hat{\lambda}_2] = \frac{\lambda^2}{n-2},$$

and from 1.(d) that the Cramer-Rao bound is given by $CRB(\lambda) = \frac{\lambda^2}{n}$, and hence that $\hat{\lambda}_2$ is unbiased but not the efficient estimator for λ.

(b) Clearly, $\text{Var}[\hat{\lambda}_2] < \text{Var}[\hat{\lambda}_{ML}]$, and since $\hat{\lambda}_2$ is also unbiased, it is preferable. The same conclusion is obviously drawn by considering the mean squared errors of the two estimators.

3. ML estimator for $\beta = \frac{1}{\lambda}$

(a)

$$f(x_1, \cdots, x_n; \beta) = \prod_{i=1}^n \frac{1}{\beta} \exp\left(-\frac{x_i}{\beta}\right) = \frac{1}{\beta^n} \exp\left(-\frac{1}{\beta}\sum_{i=1}^n x_i\right)$$

$$\ln f(x_1, \cdots, x_n; \beta) = -n\ln(\beta) - \frac{1}{\beta}\sum_{i=1}^n x_i$$

$$\frac{\partial \ln f(x_1, \cdots, x_n; \beta)}{\partial \beta} = -\frac{n}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^n x_i \overset{!}{=} 0 \qquad \Longrightarrow \beta = \frac{1}{n}\sum_{i=1}^n x_i$$

$$\frac{\partial^2 \ln f(x_1, \cdots, x_n; \beta)}{\partial \beta^2} = \frac{n}{\beta^2} - 2\frac{\sum_{i=1}^n x_i}{\beta^3} \overset{\beta>0}{\propto} n\,\beta - 2\sum_{i=1}^n x_i\Bigg|_{\beta=\frac{1}{n}\sum_{i=1}^n x_i} = -\sum_{i=1}^n x_i < 0$$

$$\Longrightarrow \quad \hat{\beta}_{ML} = \frac{1}{n}\sum_{i=1}^n x_i$$

(b) From 1.(b), we have that $\sum_{i=1}^n x_i \sim \mathcal{G}(n, \beta)$ therefore

$$\mathbb{E}[\hat{\beta}_{ML}] = \frac{1}{n} n\beta = \beta \qquad \Longrightarrow \qquad \text{unbiased}$$

$$\text{Var}[\hat{\beta}_{ML}] = \frac{1}{n^2} n\beta^2 = \frac{\beta^2}{n}.$$

(c)

$$\mathbb{E}\left[\frac{\partial^2 \ln f(x_1, \cdots, x_n; \beta)}{\partial \beta^2}\right] = \frac{n}{\beta^2} - 2\frac{\sum_{i=1}^n \mathbb{E}[x_i]}{\beta^3} = -\frac{n}{\beta^2}$$

$$\Longrightarrow \quad CRB(\beta) = \frac{\beta^2}{n}$$

$$\Longrightarrow \quad \hat{\beta}_{ML} \text{ is efficient.}$$

4. Bayesian estimation with Jeffrey's prior for λ

   (a) From 1.(d), we immediately have $p(\lambda) \propto \sqrt{I(\lambda)} \propto \frac{\sqrt{n}}{\lambda} \propto \frac{1}{\lambda}$.

   (b)

$$f(\lambda|x) \propto \frac{1}{\lambda}\lambda^n \exp\left(-\lambda\sum_{i=1}^n x_i\right) \propto \mathcal{G}\left(n, \frac{1}{\sum_{i=1}^n x_i}\right) \quad \Longrightarrow \quad \hat{\lambda}^J_{MMSE} = \frac{n}{\sum_{i=1}^n x_i} = \hat{\lambda}_{ML}.$$

$$\ln f(\lambda|x) \propto (n-1)\ln(\lambda) - \lambda\sum_{i=1}^n x_i$$

$$\frac{\partial \ln f(\lambda|x)}{\partial\lambda} \propto \frac{n-1}{\lambda} - \sum_{i=1}^n x_i \overset{!}{=} 0 \qquad \Longrightarrow \qquad \hat{\lambda}^J_{MAP} = \frac{n-1}{\sum_{i=1}^n x_i} = \hat{\lambda}_2.$$

$$\frac{\partial^2 \ln f(\lambda|x)}{\partial\lambda^2} \propto -\frac{n-1}{\lambda^2} < 0$$

5. Bayesian estimation with Gamma prior $\mathcal{G}(k,\theta)$ for λ: $\quad \lambda \sim \mathcal{G}(k,\theta)$

   (a)

$$\begin{aligned} f(\lambda|x,k,\theta) &\propto f(x_1,\cdots,x_n;\lambda)p_\lambda(\lambda;k,\theta) \\ &\propto \lambda^n \exp\left(-\lambda\sum_{i=1}^n x_i\right)\frac{1}{\Gamma(k)\theta^k}\lambda^{k-1}\exp(-\lambda/k) \\ &\propto \mathcal{G}\left(n+k, \frac{1}{\sum_{i=1}^n x_i + \frac{1}{\theta}}\right). \end{aligned}$$

   (b) Since the posterior law of λ is $\mathcal{G}\left(n+k, \frac{1}{\sum_{i=1}^n x_i + \frac{1}{\theta}}\right)$, we immediately have

$$\hat{\lambda}_{MMSE} = \frac{n+k}{\sum_{i=1}^n x_i + \frac{1}{\theta}}.$$

$$\begin{aligned} \ln f(\lambda|x,k,\theta) &\propto (n+k-1)\ln(\lambda) - \lambda\left(\sum_{i=1}^n x_i + \frac{1}{\theta}\right) \\ \frac{\partial \ln f(\lambda|x,k,\theta)}{\partial\lambda} &\propto \frac{n+k-1}{\lambda} - \sum_{i=1}^n x_i - \frac{1}{\theta} \overset{!}{=} 0 \\ \Longrightarrow \quad \hat{\lambda}_{MAP} &= \frac{n+k-1}{\sum_{i=1}^n x_i + \frac{1}{\theta}} \end{aligned}$$

   (One verifies that $\left.\frac{\partial^2 \ln f(\lambda|x,k,\theta)}{\partial\lambda^2}\right|_{\lambda=\hat{\lambda}_{MAP}} < 0$.)
   Since $\sum_{i=1}^n x_i \gg \frac{1}{\theta}$ and $n \gg k$ the estimators become equivalent as as $n \to \infty$.

Exercise 2: Detection

1. We want to test the hypotheses $H_0 : \theta = \theta_0$ and $H_1 : \theta = \theta_1$ with $\theta_1 > \theta_0$.

   (a)

$$
\begin{aligned}
\text{reject } H_0 \text{ if} \quad & \frac{f(x_1, \cdots, x_n; k, \theta|H_1)}{f(x_1, \cdots, x_n; k, \theta|H_0)} > K_\alpha \\
& \frac{\prod_{i=1}^n x_i^{k-1} \frac{k}{\theta_1} \exp\left(-\frac{x_i^k}{\theta_1}\right)}{\prod_{i=1}^n x_i^{k-1} \frac{k}{\theta_0} \exp\left(-\frac{x_i^k}{\theta_0}\right)} > K_\alpha \\
& \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left(-\sum_{i=1}^n x_i^k \left(\frac{1}{\theta_1} - \frac{1}{\theta_0}\right)\right) > K_\alpha \\
& n(\ln\theta_0 - \ln\theta_1) + \sum_{i=1}^n x_i^k \underbrace{\left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right)}_{>0} > \ln K_\alpha \\
& T = \sum_{i=1}^n x_i^k > t_\alpha = \frac{\theta_0\theta_1}{\theta_1 - \theta_0} \left(\ln K_\alpha + n(\ln\theta_1 - \ln\theta_0)\right)
\end{aligned}
$$

   (b) i. $U = T^k$ and $T = U^{\frac{1}{k}}$ and $\frac{dT}{dU} = \frac{1}{k} U^{\frac{1}{k}-1}$ and hence

$$
g(u) = f(t(u)) \left|\frac{dT}{dU}\right| = \frac{k}{\theta} u^{\frac{k-1}{k}} \exp\left(-\frac{u}{\theta}\right) \frac{1}{k} u^{\frac{1}{k}-1} = \frac{1}{\theta} \exp\left(-\frac{u}{\theta}\right)
$$

   and hence $U \sim \mathcal{G}(1, \theta)$.

   ii. Since $\varphi_U(t) = \mathbb{E}\left[e^{itU}\right] = (1 - it\theta)^{-\beta}$ with $\beta = 1$, it follows that

$$
\varphi_{T=\sum_{m=1}^n U_m}(t) = \mathbb{E}\left[e^{it\sum_{m=1}^n U_m}\right] = \prod_{m=1}^n \mathbb{E}\left[e^{itU_m}\right] = (1 - it\theta)^n
$$

   and hence $T \sim \mathcal{G}(n, \theta)$.

   Therefore, under H_0, $T \sim \mathcal{G}(n, \theta_0)$ and under H_1, $T \sim \mathcal{G}(n, \theta_1)$.

   (c)

$$
\begin{aligned}
\alpha &= P[\text{reject } H_0 | H_0 \text{ true}] = P[T > t_\alpha | T \sim \mathcal{G}(n, \theta_0)] \\
&= 1 - \int_0^{t_\alpha} \frac{x^{n-1}}{\Gamma(n)\theta_0^n} \exp(-x/\theta_0) dx = 1 - F_0(t_\alpha) \\
&\Longrightarrow t_\alpha = F_0^{-1}(1 - \alpha)
\end{aligned}
$$

   (d)

$$
\begin{aligned}
\pi &= P[\text{reject } H_0 | H_1 \text{ true}] = P[T > t_\alpha | T \sim \mathcal{G}(n, \theta_1)] \\
&= 1 - \int_0^{t_\alpha} \frac{x^{n-1}}{\Gamma(n)\theta_1^n} \exp(-x/\theta_1) dx = 1 - F_1(t_\alpha) \\
&= 1 - F_1(F_0^{-1}(1 - \alpha))
\end{aligned}
$$

2. We first need to determine the maximum likelihood estimator for θ,

$$
\begin{aligned}
f(x_1,\cdots,x_n;\beta,\theta) &= \prod_{i=1}^{n} k\frac{x_i^{k-1}}{\theta}\exp\left(-\frac{x_i^k}{\theta}\right) = \frac{k^n}{\theta^n}\exp\left(-\frac{\sum_{i=1}^n x_i^k}{\theta}\right)\prod_{i=1}^{n} x_i^k \\
\ln f(x_1,\cdots,x_n;\beta,\theta) &= n\ln(k) - n\ln(\theta) - \frac{1}{\theta}\sum_{i=1}^{n} x_i^k + \ln\left(\prod_{i=1}^{n} x_i^k\right) \\
\frac{\partial \ln f(x_1,\cdots,x_n;\beta,\theta)}{\partial\theta} &= -\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i^k}{\theta^2} \stackrel{!}{=} 0 \qquad \Longrightarrow \theta = \frac{1}{n}\sum_{i=1}^{n} x_i^k \\
\left.\frac{\partial^2 \ln f(x_1,\cdots,x_n;\beta,\theta)}{\partial\theta^2}\right|_{\theta=\frac{1}{n}\sum_{i=1}^n x_i^k} &= -\sum_{i=1}^{n} x_i^k < 0 \quad \Longrightarrow \quad \hat{\theta}_{ML} = \frac{1}{n}\sum_{i=1}^{n} x_i^k.
\end{aligned}
$$

The generalized likelihood ratio test is given by

$$
\begin{aligned}
\text{reject } H_0 \text{ if} \quad & \frac{\sup_{\theta\in\Omega_1} f(x_1,\cdots,x_n|\theta)}{\sup_{\theta\in\Omega_0} f(x_1,\cdots,x_n|\theta)} = \frac{f(x_1,\cdots,x_n|\theta=\hat{\theta}_{ML(1)})}{f(x_1,\cdots,x_n|\theta=\theta_0)} > K_\alpha \\
& \frac{\prod_{i=1}^n k\frac{x_i^{k-1}}{\hat{\theta}_{ML(1)}}\exp\left(-\frac{x_i^k}{\hat{\theta}_{ML(1)}}\right)}{\prod_{i=1}^n k\frac{x_i^{k-1}}{\theta_0}\exp\left(-\frac{x_i^k}{\theta_0}\right)} > K_\alpha \\
& \left(\frac{n\theta_0}{\sum_{i=1}^n x_i^k}\right)^n \exp\left(-\left(\frac{n}{\sum_{i=1}^n x_i^k} - \frac{1}{\theta_0}\right)\sum_{i=1}^{n} x_i^k\right) > K_\alpha \\
& n\ln(n\theta_0) - n\ln\left(\sum_{i=1}^{n} x_i^k\right) - \left(\frac{n}{\sum_{i=1}^n x_i^k} - \frac{1}{\theta_0}\right)\sum_{i=1}^{n} x_i^k > \ln K_\alpha \\
& T = -\ln\left(\sum_{i=1}^{n} x_i^k\right) + \frac{1}{n\theta_0}\sum_{i=1}^{n} x_i^k \quad > \quad t_\alpha = \frac{\ln K_\alpha}{n} + 1 - \ln(n) - \ln(\theta_0).
\end{aligned}
$$

3. (a) Since the parameter θ is unknown, the Kolmogorov test can not be applied. The χ^2 test is appropriate for this problem.

   (b)

$$
\begin{aligned}
f(t_1,\cdots,t_M;n,\theta) &= \frac{\prod_{m=1}^M t_m^{n-1}}{\Gamma(n)^M\theta^{nM}}\exp\left(-\frac{\sum_{m=1}^M t_m}{\theta}\right) \\
\ln f(t_1,\cdots,t_M;n,\theta) &\propto -nM\ln(\theta) - \frac{1}{\theta}\sum_{m=1}^{M} t_m \\
\frac{\partial \ln f(t_1,\cdots,t_M;n,\theta)}{\partial\theta} &= -\frac{nM}{\theta} + \frac{\sum_{m=1}^M t_m}{\theta^2} \stackrel{!}{=} 0 \qquad \Longrightarrow \hat{\theta}_{ML} = \frac{\sum_{m=1}^M t_m}{nM}
\end{aligned}
$$

   and hence

$$
\hat{\theta}_{ML} = \frac{\sum_{m=1}^M t_m}{nM} = \frac{1250}{625} = 2
$$

   The mean is given by $\mu = n\hat{\theta}_{ML} = 50$ and the standard deviation by $\sigma = \sqrt{n\hat{\theta}_{ML}^2} = 10$.

   (c) With $a = F^{-1}(3/4)$, the classes for testing for the standard Normal distribution $\mathcal{N}(0,1)$ are given by

$$
(-\infty,-a],\ (-a,0]\ (0,a],\ (a,\infty)
$$

and the classes for $\mathcal{N}(\mu = 50, \sigma^2 = 10^2)$ are therefore

$$\begin{aligned}
C_1 &= (-\infty, \mu - \sigma a] = (-\infty, 43.25] \\
C_2 &= (\mu - \sigma a, \mu] = (43.25, 50] \\
C_3 &= (\mu, \mu + \sigma a] = (50, 56.75] \\
C_4 &= (\mu + \sigma a, \infty] = (56.75, \infty]
\end{aligned}$$

(d) The test statistic is given by

$$\phi = \sum_{k=1}^{K} \frac{(n_k - Mp_k)^2}{Mp_k}$$

with $K = 4$, $p_k = 1/4$ and $(n_1, n_2, n_3, n_4) = (7, 5, 6, 7)$. Hence,

$$\phi = \frac{4}{25}\left(\frac{9}{16} + \frac{25}{16} + \frac{1}{16} + \frac{9}{16}\right) = \frac{11}{25} = 0.44.$$

Since the unknown parameter θ was replaced with its maximum likelihood estimate, the critical value of the test is given by the quantile of the χ^2 distribution with $N = K - 1\ - 1$ degrees of freedom, $t_\alpha = 4.61 > 0.44$. The test with significance $\alpha = 0.1$ therefore does not reject the hypothesis that the sample is Normal.