Tutorial letter 201/1/2018

Statistical Inference II

STA2602

Semester 1

Department of Statistics

Solutions to Assignment 1

QUESTION 1 **[25]**

(a) **(6)**

$$\begin{aligned} MSE(\hat{\theta}) &= E[(\hat{\theta}-\theta)^2] \\ &= E[[(\hat{\theta}-E(\hat{\theta}))+(E(\hat{\theta})-\theta)]^2] \\ &= E[(\hat{\theta}-E(\hat{\theta}))^2+2(\hat{\theta}-E(\hat{\theta}))(E(\hat{\theta})-\theta)+(E(\hat{\theta})-\theta))^2] \\ &= E[(\hat{\theta}-E(\hat{\theta}))^2]+2(E(\hat{\theta})-E(\hat{\theta}))(E(\hat{\theta})-\theta)+(E(\hat{\theta})-\theta))^2 \\ &= E[(\hat{\theta}-E(\hat{\theta}))^2]+0+(E(\hat{\theta})-\theta))^2 \\ &= V(\hat{\theta})+[B(\hat{\theta})]^2 \end{aligned}$$

(b) **(19)**

(i) $E(\hat{\theta}_3)=\alpha E(\hat{\theta}_1)+(1-\alpha)E(\hat{\theta}_2)=\alpha\theta+(1-\alpha)\theta=\theta.$ (3)

(ii) $V(\hat{\theta}_3)=\alpha^2V(\hat{\theta}_1)+(1-\alpha)^2V(\hat{\theta}_2)=[\alpha^2+(1-\alpha)^2]\sigma^2=(2\alpha^2-2\alpha+1)\sigma^2.$ (4)

(iii) We find the value of α which solves the equation

$$0=\frac{\partial V(\hat{\theta}_3)}{\partial\alpha}=(4\alpha-2)\sigma^2.$$

The solution is $\alpha=\frac{1}{2}$ minimizes $V(\hat{\theta}_3)$. (5)

(iv) $V(\hat{\theta}_3)=\alpha^2V(\hat{\theta}_1)+(1-\alpha)^2V(\hat{\theta}_2)=\alpha^2\sigma^2+2(1-\alpha)^2\sigma^2=[3\alpha^2-4\alpha+2]\sigma^2.$ (3)
Hence the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_3$ is

$$eff(\hat{\theta}_1,\hat{\theta}_3)=\frac{V(\hat{\theta}_1)}{V(\hat{\theta}_3)}=\frac{1}{3\alpha^2-4\alpha+2}.$$

(2)

(v) $eff(\hat{\theta}_1,\hat{\theta}_3)\Big|_{\alpha=1/2}=\frac{1}{3\alpha^2-4\alpha+2}\Big|_{\alpha=1/2}=\frac{4}{3}.$ (2)

QUESTION 2 **[8]**

$E(\hat{\theta}_n)=\left(\frac{n+1}{n}\right)E(Y_{(n)})=\left(\frac{n+1}{n}\right)\times\left(\frac{n}{n+1}\right)\theta=\theta$. Hence $\lim_{n\to\infty}E(\hat{\theta}_n)=\theta.$ (3)

$V(\hat{\theta}_n)=\left(\frac{n+1}{n}\right)^2V(Y_{(n)})=\left(\frac{n+1}{n}\right)^2\times\left(\frac{n}{(n+2)(n+1)^2}\right)\theta^2=\frac{\theta^2}{n(n+2)}$ and
$\lim_{n\to\infty}V(\hat{\theta}_n)=0.$ (5)

That $\lim_{n\to\infty}E(\hat{\theta}_n)=\theta$ and $\lim_{n\to\infty}V(\hat{\theta}_n)=0$ implies that $\hat{\theta}_n=\left(\frac{n+1}{n}\right)Y_{(n)}$ is a consistent estimator of θ.

QUESTION 3 **[7]**

(a) **(3)**

$$\begin{aligned} L(\beta) &= \prod_{i=1}^{n} f(x_i, \beta) = \prod_{i=1}^{n} \beta e^{-\beta x_i} \\ &= \beta e^{-\beta x_1} \times \beta e^{-\beta x_2} \times ... \times \beta e^{-\beta x_n} \\ &= \beta^n e^{-\beta \sum_{i=1}^{n} x_i} \end{aligned}$$

(b) **(4)**

From part (a) we have $L(\beta) = \beta^n e^{-\beta \sum_{i=1}^{n} x_i} = g\left(\sum_{i=1}^{n} x_i, \beta\right) \times h(x_1, x_2, ..., x_n)$ where $g\left(\sum_{i=1}^{n} x_i, \beta\right) = \beta^n e^{-\beta \sum_{i=1}^{n} x_i}$ (**depends on the sample only through** $\sum_{i=1}^{n} x_i$) and $h(x_1, x_2, ..., x_n = 1$ (**independent of** β). This means, by the factorization criterion, $\sum_{i=1}^{n} X_i$ is a sufficient statistic for β.

QUESTION 4 **[10]**

(a) **(3)**

$E(X_i) = \dfrac{1}{\beta}$ means $\beta = \dfrac{1}{E(X_i)}$. The method-of-moments estimator is $\tilde{\beta}$ which is obtained by replacing $E(X_i)$ in $\beta = \dfrac{1}{E(X_i)}$ with a corresponding sample moment which in this case is $\bar{X}$. Hence $\tilde{\beta} = \dfrac{1}{\bar{X}}$ is the method-of-moments estimator of β.

(b) **(7)**

From **QUESTION 3** the log-likelihood function is

$$l(\beta) = \ln L(\beta) = n \ln \beta - \beta \sum_{i=1}^{n} x_i = n \ln \beta - n\beta\bar{x}$$

where $\bar{x} = \dfrac{1}{n} \sum_{i=1}^{n} x_i$. (2)

The maximum likelihood estimate of β is $\hat{\beta}$ which solves the equation

$$0 = l'(\hat{\beta}) = \left.\frac{\partial l(\beta)}{\partial \beta}\right|_{\beta=\hat{\beta}} = \frac{n}{\hat{\beta}} - n\bar{x}.$$

(3)

The solution is $\hat{\beta} = \dfrac{1}{\bar{x}}$. Hence the maximum likelihood estimator of β is $\hat{\beta} = \dfrac{1}{\bar{X}}$. (2)

QUESTION 5 **[20]**

(a) **(6)**

$E(Y_i) = \beta_0 + E(\epsilon_i) = \beta_0.$ (2)

The method-of-moments estimator is $\tilde{\beta}$ which is obtained by replacing $E(Y_i)$ in $E(Y_i) = \beta_0$ with a corresponding sample moment which in this case is $\bar{Y}$. Hence $\tilde{\beta}_0 = \bar{Y}$ is the method-of-moments estimator of β_0. (3)

$V(\tilde{\beta}_0) = V(\bar{Y}) = \frac{\sigma^2}{n}.$ (1)

(b) **(7)**

The likelihood of the sample is

$$\begin{aligned} L(\beta_0) &= \prod_{i=1}^{n} f(y_i, \beta_0) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y_i-\beta_0)^2} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\beta_0)^2} \end{aligned}$$

(2)

and the log-likelihood is

$$l(\beta_0) = \ln L(\beta_0) = -n/2 \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0)^2.$$

(1)

The maximum likelihood estimate of β_0 is $\hat{\beta}_0$ which solves the equation

$$0 = l'(\hat{\beta}_0) = \left.\frac{\partial l(\beta_0)}{\partial \beta_0}\right|_{\beta_0=\hat{\beta}_0} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \hat{\beta}_0).$$

(2)

The solution is $\hat{\beta}_0 = \bar{y}$. Hence the maximum likelihood estimator of β_0 is $\hat{\beta}_0 = \bar{Y}$. (2)

(c) **(7)**

The least squares estimator of β_0 is $\hat{\beta}_0$ which solves the equation

$$0 = \left.\frac{\partial SSE}{\partial \beta_0}\right|_{\beta_0=\hat{\beta}_0} = -2\sum_{i=1}^{n}(y_i - \hat{\beta}_0).$$

(4)

The solution is $\hat{\beta}_0 = \bar{y}$. Hence the least squares estimator of β_0 is $\hat{\beta}_0 = \bar{Y}$ which is the same as the maximum likelihood estimator of β_0. (3)

QUESTION 6 **[30]**

(a) **(6)**

$$S_{xx}=\sum_{i=1}^{5}x_i^2-\frac{1}{5}\left(\sum_{i=1}^{5}x_i\right)^2=55-\frac{(15)^2}{5}=10,$$ (2)

$$S_{yy}=\sum_{i=1}^{5}y_i^2-\frac{1}{5}\left(\sum_{i=1}^{5}y_i\right)^2=26-\frac{(10)^2}{5}=6,$$ (2)

and $S_{xy}=\sum_{i=1}^{5}x_iy_i-\frac{1}{5}\left(\sum_{i=1}^{5}x_i\right)\left(\sum_{i=1}^{5}y_i\right)=37-\frac{(15)(10}{5}=7.$ (2)

(b) **(5)**

$$r=\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}=\frac{7}{\sqrt{(10)(6)}}=0.9034.$$ (3)

The correlation coefficient is positive and large (**close to 1**). This means there is a strong positive linear relationship between advertising expenditure (x) and sales revenue (y).

(2)

(c) **(6)**

$$\hat{\beta}_1=\frac{S_{xy}}{S_{xx}}=\frac{7}{10}=0.7,$$ (3)

and $\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}=\frac{10}{5}-(0.7)\frac{15}{5}=-0.1.$ (3)

(**The fitted model is** $\hat{y}=-0.1+0.7x$.)

(d) **(3)**

$$SSE=S_{yy}-\hat{\beta}_1S_{xy}=6-(0.7)(7)=1.1.$$ (2)

Hence the estimate of σ^2 is $s^2=\frac{SSE}{5-2}=\frac{1.1}{3}=0.3667.$ (1)

(e) **(2)**

The estimate of $V(\hat{\beta}_1)$ is $s^2_{\hat{\beta}_1}=\frac{s^2}{S_{xx}}=\frac{0.3667}{10}=0.0367.$

(f) **(8)**

The test statistic is $t=\frac{\hat{\beta}_1}{s_{\hat{\beta}_1}}=\frac{0.7}{0.1916}=3.654.$ (3)

The decision rule is "Reject H_0 if $|t|>t_{0.025}=3.182$ (**obtained from the** $t-tables$ **using 3 degrees of freedom**)" or "Reject H_0 if either $t<-t_{0.025}=-3.182$ or $t>t_{0.025}=3.182$.

(3)

Since $t=3.654>t_{0.025}=3.182$ we reject H_0 and conclude, at the 0.05 level of significance, that there is a strong linear between advertising expenditure (x) and sales revenue (y). (2)

TOTAL: [100]