200B HW#5 solution

7.8 Jointly Sufficient Statistics

6. The joint p.d.f. is

$$f_n(\mathbf{x}|\theta) = \left\{ \prod_{j=1}^{n} b(X_j) \right\} [a(\theta)]^n \exp \left\{ \sum_{i=1}^{k} c_i(\theta) \sum_{j=1}^{k} d_i(X_j) \right\}.$$

It follows from the factorization theorem with $u(\mathbf{x}) = \prod_{j=1}^{n} b(X_j)$ and $v((T_1, \ldots, T_k), \theta) = [a(\theta)]^n \exp \left\{ \sum_{i=1}^{k} c_i(\theta) T_i \right\}$ that $T_1, \ldots, T_k$ are jointly sufficient statistics for θ.

10. The p.d.f. of the uniform distribution is $f(x|\theta) = \frac{1}{\theta} \mathbf{1}_{\{x \in [0,\theta]\}}$, so the likelihood is

$$\tilde{L}(\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(X_i|\theta) = \frac{1}{\theta^n} \mathbf{1}_{\{X_{(n)} \leq \theta\}} \mathbf{1}_{\{X_{(1)} \geq 0\}}$$

It implies that the M.L.E. $\hat{\theta} = X_{(n)}$ since the likelihood function is decreasing in θ.

$T = \hat{\theta}$ is the sufficient statistic by the factorization theorem with $u(\mathbf{x}) = \mathbf{1}_{\{X_{(1)} \geq 0\}}$ and $v(T, \theta) = \frac{1}{\theta^n} \mathbf{1}_{\{T \leq \theta\}}$, so it is a minimal sufficient statistic by Corollary in page 30 of the lecture notes.

12. The likelihood is

$$\tilde{L}(\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(X_i|\theta) = \frac{2^n \prod_{i=1}^{n} X_i}{\theta^{2n}} \mathbf{1}_{\{X_{(n)} \leq \theta\}} \mathbf{1}_{\{X_{(1)} \geq 0\}}.$$

The M.L.E. of θ is $\hat{\theta} = X_{(n)}$ since the likelihood function is decreasing in θ. $T = X_{(n)}$ is also a sufficient statistic by the factorization theorem with $u(\mathbf{x}) = 2^n (\prod_{i=1}^{n} X_i) \mathbf{1}_{\{X_{(1)} \geq 0\}}$ and $v(T, \theta) = \frac{\mathbf{1}_{\{T \leq \theta\}}}{\theta^{2n}}$.

The median of the distribution is the value m such that $\int_0^m f(x|\theta)dx = \int_0^m x/\theta^2 dx = \theta^{-2} x^2 |_0^m = m^2/\theta^2 = 1/2$, it implies that $m = \theta/\sqrt{2}$.

By the invariance property of M.L.E., $\hat{m} = \hat{\theta}/\sqrt{2}$ is the M.L.E. of m. Note that $\hat{m}$ is also a sufficient statistic, being a one to one function of the sufficient statistics, $T = \hat{\theta}$. So by Corollary in page 30 of the lecture notes it is a minimal sufficient statistic.

16. Let the prior distribution be Gamma(α, β). Then $\xi(\lambda) \propto \lambda^{\alpha-1}e^{-\beta\lambda}$. The joint pmf of data is $f_n(\mathbf{x}|\lambda) = \frac{\lambda^{\sum_{i=1}^n X_i} e^{-n\lambda}}{(\prod_{i=1}^n X_i!)} \propto \lambda^{\sum_{i=1}^n X_i} e^{-n\lambda}$. So the posterior $\xi(\lambda|\mathbf{x}) \propto \xi(\lambda) f_n(\mathbf{x}|\lambda) \propto \lambda^{\alpha+\sum_{i=1}^n X_i - 1} e^{-(n+\beta)\lambda}$, which is identified as the density of Gamma$(\alpha + \sum_{i=1}^n x_i, \beta + n)$ distribution. It follows from Theorem 7.3.2 that the Bayes estimator of λ under the square error loss is the posterior mean, given by, $\hat{\lambda} = (\alpha + \sum_{i=1}^n X_i)/(\beta + n)$. $T = \sum_{i=1}^n X_i$ is a sufficient statistic for λ by the factorization theorem with $u(\mathbf{x}) = \frac{1}{(\prod_{i=1}^n X_i!)}$ and $v(T, \lambda) = \lambda^T e^{-n\lambda}$. The Bayes estimator $\hat{\lambda}$ is also a sufficient statistic for λ since it is a one-to-one function of $\sum_{i=1}^n X_i$. Hence, the Bayes estimator $\hat{\lambda}$ is a minimal sufficient statistic by the Corollary in page 30 of the lecture notes.

7.9 Improving an Estimator

2. It is derived that $T = X_{(n)}$ is a sufficient statistic in the lecture notes page 24. Since $2\bar{X}_n$ is not a function of the sufficient statistic, it is improvable by Rao-Blackwell theorem ($\delta_0(T) = E[2\bar{X}_n|T]$ has smaller MSE than $2\bar{X}_n$). So there exist an estimator $\delta_0(T)$ that dominates the estimator $\delta(\mathbf{X}) = 2\bar{X}_n$ in terms of smaller MSE and therefore $2\bar{X}_n$ is inadmissible. (See page 31-32 in the lecture notes).

6. The likelihood is

$$f(\mathbf{x}|\alpha) = \frac{\beta^{n\alpha}}{[\Gamma(\alpha)]^n} \left(\prod_{i=1}^n X_i\right)^{\alpha-1} \exp\left(-\beta \sum_{i=1}^n X_i\right) \mathbf{1}_{\{X_{(1)}>0\}}.$$

$T = \prod_{i=1}^n X_i$ is a sufficient statistic by the factorization theorem with $u(\mathbf{x}) = \exp\left(-\beta \sum_{i=1}^n X_i\right) \mathbf{1}_{\{X_{(1)}>0\}}$ and $v(T, \alpha) = \frac{\beta^{n\alpha}}{[\Gamma(\alpha)]^n} T^{\alpha-1}$. The Rao-Blackwell theorem says that when using the squared error loss function, an estimator which is not a function of a sufficient statistic can be improved. Since $\bar{X}_n$ is not a function of the sufficient statistic $\prod_{i=1}^n X_i$, it is inadmissible.(See page 31-32 in the lecture notes).

7.10 Supplementary Exercises

14. The joint p.d.f. of $X_1, \dots, X_n$ can be written in the form

$$f_n(\mathbf{x}|\beta,\theta) = \beta^n \exp\left(n\beta\theta - \beta\sum_{i=1}^{n} X_i\right)$$

for $X_{(1)} \geq \theta$, and $f_n(\mathbf{x}|\beta,\theta) = 0$ otherwise. Using indicator function, the p.d.f. is

$$f_n(\mathbf{x}|\beta,\theta) = \beta^n \exp\left(n\beta\theta - \beta\sum_{i=1}^{n} X_i\right)\mathbf{1}_{\{X_{(1)}\geq\theta\}}.$$

Hence, $(T_1, T_2) = (\sum_{i=1}^{n} X_i, X_{(1)})$ is a pair of jointly sufficient statistics by the factorization theorem with $u(\mathbf{x}) = 1$ and $v((T_1,T_2),(\beta,\theta)) = \beta^n \exp\left(n\beta\theta - \beta T_1\right)\mathbf{1}_{\{T_2\geq\theta\}}$.

17. The joint pdf is

$$f_n(\mathbf{x}|x_0,\alpha) = \frac{\alpha^n x_0^{n\alpha}}{\{\prod_{i=1}^{n} X_i\}^{\alpha+1}}, \text{ for } X_{(1)} \geq x_0.$$

It is an increasing function of x_0 so that $\hat{x}_0 = X_{(1)}$ is the M.L.E. of x_0. If we substitute $\hat{x}_0$ for x_0 and let $L(\alpha)$ denote the logarithm of the resulting likelihood function, then

$$L(\alpha) = n\log\alpha + n\alpha\log\hat{x}_0 - (\alpha+1)\sum_{i=1}^{n}\log x_i$$

,

$$\frac{dL(\alpha)}{d\alpha} = \frac{n}{\alpha} + n\log\hat{x}_0 - \sum_{i=1}^{n}\log x_i$$

and

$$\frac{dL(\alpha)^2}{d^2\alpha} = -\frac{n}{\alpha^2} < 0$$

.

Hence, by setting $\frac{dL(\alpha)}{d\alpha}$ equal to 0, we find that

$$\hat{\alpha} = \left(\frac{1}{n}\sum_{i=1}^{n}\log x_i - \log\hat{x}_0\right)^{-1}.$$

As long as $\hat{\alpha} > 0$, it is the M.L.E of α. If $\hat{\alpha} \leq 0$, then $L(\alpha)$ is an increasing function of α and the M.L.E. is given by $\hat{\alpha} = 0$.

18. Using indicator function, the likelihood is

$$f_n(\mathbf{x}|x_0,\alpha) = \frac{\alpha^n x_0^{n\alpha}}{(\prod_{i=1}^{n} X_i)^{\alpha+1}}\mathbf{1}_{\{X_{(1)}\geq x_0\}}.$$

$T_1 = X_{(1)}$ and $T_2 = \prod_{i=1}^n X_i$ are jointly sufficient statistics for x_0 and α by factorization theorem with $u(\mathbf{x}) = 1$ and $v((T_1, T_2), (x_0, \alpha)) = \frac{\alpha^n x_0^{n\alpha}}{T_2^{\alpha+1}} \mathbf{1}_{\{T_1 \geq x_0\}}$. Here $(\hat{x}_0, \hat{\alpha})$ form a one-to-one transform of (T_1, T_2), so $\hat{x}_0$ and $\hat{\alpha}$ are also jointly sufficient statistics. Note that they are M.L.E. for x_0 and α by the last question. Hence, they are minimal jointly sufficient statistics by Corollary in page 30 of the lecture notes.