CHALMERS | UNIVERSITY OF GOTHENBURG

Annika Lang

Statistical Inference Principles – Spring 2018

Assignment 2

1. Let X be a random sample of size n and $X_1 \sim \mathcal{N}(\mu, \sigma^2)$. Assume that σ^2 is known. Show that the sample mean $T(X) = \bar{X}$ is a sufficient statistic for the parameter μ.

2. Let X be a random sample of size n with PMDF $f_{X_1}(\cdot|\theta)$ from an exponential family given by

$$f_{X_1}(x|\theta) := h(x)c(\theta)\exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right),$$

where $\theta := (\theta_1, \ldots, \theta_d)$, $d \leq k$. Show that

$$T(X) := \left(\sum_{j=1}^{n} t_1(X_j), \ldots, \sum_{j=1}^{n} t_k(X_j)\right)$$

is a sufficient statistic for θ.

3. Let (X_1, X_2) be a random sample of the discrete distribution given by

$$P(X_1 = \theta) = P(X_1 = \theta + 1) = P(X_1 = \theta + 2) = \frac{1}{3}.$$

Set $R := X_{(2)} - X_{(1)}$ and $M := (X_{(1)} + X_{(2)})/2$.

(a) Show that (R, M) is a minimal sufficient statistic for θ.

(b) Furthermore, show that R is an ancillary statistic.

4. Let X be a random sample of size n with PMDF f_X and $X_{(\cdot)}$ its order statistic. Then we have shown in the lecture that for $j = 1, \ldots, n$ and $x \in \mathbb{R}$

$$F_{X_{(j)}}(x) = \sum_{k=j}^{n} \binom{n}{k} F_{X_1}(x)^k (1 - F_{X_1}(x))^{n-k}.$$

Please turn!

a) Assume that f_X is a PMF with $f_{X_1}(x_i) = p_i > 0$, where $x_1 < x_2 < \cdots$ and $x_i \in X_1(\Omega)$. Show that for $j = 1, \ldots, n$

$$f_{X_{(j)}}(x_i) = \sum_{k=j}^{n} \binom{n}{k} \left(F_{X_1}(x_i)^k (1 - F_{X_1}(x_i))^{n-k} - F_{X_1}(x_{i-1})^k (1 - F_{X_1}(x_{i-1}))^{n-k} \right).$$

b) Assume that f_X is a PDF. Show that the PDF of the order statistic is then given by

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_{X_1}(x) F_{X_1}(x)^{j-1} (1 - F_{X_1}(x))^{n-j}$$

for $j = 1, \ldots, n$ and $x \in X_1(\Omega) \subset \mathbb{R}$.

Deadline: Thursday, February 1, 2018, send an email before 14.30 with a list of solved problems.
Webpage: `http://www.math.chalmers.se/Stat/Grundutb/GU/MSF100/S18/`
Requirement: 75% of the exercises solved, two presentations in the exercise class