Exam STA3000, 2020

Instructions: Please write your answers on this exam in the spaces provided.

Name:

1. Suppose that a response variable is generated by randomly selecting an individual ω from a population Ω and obtaining the measurement $x = X(\omega) \in \{1, 2, 3\}$. The proportions of elements of Ω taking these values is given by one of three possibilities, that we label by a, b and c, as provided in the following table.

	$x = 1$	$x = 2$	$x = 3$
a	1/3	1/3	1/3
b	1/6	1/6	2/3
c	1/4	1/4	1/2

Suppose we observe a single measurement x.

(a) (5 marks) Identify the statistical model, namely, provide the sample space $\mathcal{X}$, parameter space Θ and family of distributions f_θ.

(b) (5 marks) Provide the sampling distributions of $\theta_{MLE}(x)$.

(c) (5 marks) Determine whether or not θ_{MLE} is a sufficient statistic and if so, whether or not it is minimal sufficient.

(d) (5 marks) Suppose that the quantity of interest is $\Psi(\theta) = \theta$ and a loss function is given by $L(a, \psi) = 1 - I_{\{a\}}(\psi)$ and $L(\theta, \psi) = 2(1 - I_{\{\theta\}}(\psi))$ otherwise. Determine the risk function for θ_{MLE}.

(e) (5 marks) If a uniform prior is placed on Θ, then determine the prior risk of θ_{MLE}.

(f) (5 marks) Suppose δ is a decision function such that $\delta(1, \cdot)$ is degenerate at $b, \delta(2, \cdot)$ is uniform on $\{a, b\}$ and $\delta(3, \cdot)$ is degenerate at b. Determine the risk function of δ and if either δ or θ_{MLE} is preferred.

(g) (5 marks) Does δ in (f) depend on the data only through a minimal sufficient statistic? If not, modify δ so that it does and determine the risk function of this modified estimator.

2. Suppose that $x_1, \ldots, x_n$ is a sample from a $N(\mu, \sigma^2)$ distribution where $\mu \in R^1, \sigma^2 > 0$ are unknown. Suppose we wish to estimate the third quartile $\psi = \Psi(\mu, \sigma^2) = \mu + \sigma z_{0.75}$ and we use squared error loss.

(a) (5 marks) Justify why we need only consider nonrandomized estimators that are functions of $(\bar{x}, s^2)$.

(b) (5 marks) Determine an optimal unbiased estimator of ψ. Is this estimator unique?

(c) (5 marks) Is the estimator determined in (b) admssible? Hint: consider all estimators of the form $x + cs$ for some constant c.

3. (a) (5 marks) Suppose that X is a random variable and m is a median of the distribution of X. Prove that for all constants a_1, a_2 satisfying either $m \leq a_1 \leq a_2$ or $m \geq a_1 \geq a_2$ then $E(|X - a_1|) \leq E(|X - a_2|)$. Hint: write $E(|X - a_i|)$ without absolute value and show $E(|X - a_2|) - E(|X - a_1|) \geq 0$.

(b) (5 marks) Suppose in a statistical problem we want to estimate $\Psi(\theta) \in R^1$ where the range of Ψ is an interval and the loss function is $L(\theta, \psi) = |\psi - \Psi(\theta)|$. Prove that we can restrict to nonrandomized estimators. Using (a) show that a nonrandomized estimator d is unbiased for $\Psi(\theta)$ iff $\Psi(\theta)$ is a median of the distribution of d for every θ.

4. Suppose that f is a density on $\mathcal{X} = R^n$ with respect to volume measure and the statistical model is given by the set of density functions $\left\{\theta^{-n} f\left(x/\theta\right) : \theta \in \Theta\right\}$ where $\Theta = \{\theta : \theta > 0\}$.

(a) (2 marks) Show that we can consider $G = \Theta$ as a group with product $\theta_1 \cdot \theta_2 = \theta_1\theta_2$.

(b) (5 marks) Prove that $T_\theta x = \theta x$ defines an action of G on $\mathcal{X}$, that G is a symmetry group of the model and determine $\bar{T}_\theta$. Is G transitive on Θ?

(c) (3 marks) Determine the set which must be deleted from the sample space so that G acts freely on $\mathcal{X}$. Determine the orbits on $\mathcal{X}$.

(d) (3 marks) Suppose that we want to estimate $\Psi(\theta) = \theta$ and the loss function is given by $L(\theta, \psi) = \rho(\psi/\theta)$ for some convex function $\rho : (0, \infty) \to R^1$. What is the purpose of restricting to convex ρ? What is the action of the group on Ψ? Prove that G leaves this decision problem invariant. Explain what it means for an estimator to be equivariant in this problem.

(e) (4 marks) Show that $[x] = \|x\|$ is equivariant and determine a maximal invariant $D(x)$.

(f) (3 marks) Putting $s = \|x\|$ and $u = D(x)$, show that $J(x \to s, u) = s^{n-1}h(u)$ for some function h.

(g) (2 marks) Show how you would compute the Pitman estimate of θ.

(h) (5 marks) Determine the Pitman estimate when f corresponds to the $N_n(0, I)$ distribution and $\rho(\tau) = (\tau - 1)^2$.

(i) (3 marks) Explan why there isn't a uniformly most powerful invariant (under G) size α test of $H_0 : \theta = \theta_0$ versus $H_0 : \theta \neq \theta_0$.

(j) (5 marks) Under the same distribution assumption in (h), determine the uniformly most powerful unbiased size α test of $H_0 : \theta = \theta_0$ versus $H_0 : \theta \neq \theta_0$.

(k) (5 marks) Under the same distribution assumption in (h), determine the uniformly most accurate unbiased γ-confidence interval for θ.