QUAL EXAM: MATH 281 A&B

PROFESSOR JELENA BRADIC

You can use your textbooks as references for this Qual Exam. Any result (theorem, lemma) used needs to be fully quoted with the number of the result clearly expressed. Define all notations used. Specify constants if need be. Claims of any property of random variables need to be proven with details. This part of the exam should take you up to 2h to complete.

1. Problem 1. (Kernel Density Estimator)

Let $X_1, X_2, \ldots, X_n$ be i.i.d. samples drawn according to some (unknown) density ϕ on the real line. A kernel density estimator of the density function ϕ is defined as

$$\phi_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right).$$

In the above $h_n > 0$ is a smoothing parameter and K is a nonnegative function with integral one, $\int K = 1$. The performance of the estimate is typically measured by the l_1 error

$$Z(n) = f(X_1, \cdots, X_n) = \int |\phi(x) - \phi_n(x)| dx.$$

Show (prove) and identify minimal conditions (the weakest possible) under which

$$P\left(\left|\frac{Z(n)}{E[Z(n)]} - 1\right| \geq \epsilon\right) \leq \frac{\mathrm{Var}(Z(n))}{\epsilon^2 \{E[Z(n)]\}^2} \to 0.$$

2. Problem 2. (Change point detection)

Suppose that we have i.i.d. data $\{X_i = (Z_i, Y_i)\}$ for $i = 1, \ldots, n$. Here let $Z_i \sim \mathrm{Unif}(0, 1)$ and

$$Y_i = 1\{0 \leq Z_i \leq \theta_0\} + \epsilon_i.$$

Let ϵ_i be i.i.d. $\mathcal{N}(0, 1)$ and let them be independent of Z_i. The goal is to estimate the unknown parameter $\theta_0 \in (0, 1)$. Show that

$$\hat{\theta}_n \xrightarrow{p} \theta_0.$$

Here,

$$\hat{\theta}_n = \arg\min_{\theta \in [0,1]} n^{-1} \sum_{i=1}^{n} (Y_i - 1\{0 \leq Z_i \leq \theta\})^2$$

Moreover, find the rate of convergence of $\hat{\theta}_n$. Then, generalize this estimator to the setting where

$$Y_i = \alpha 1\{0 \leq Z_i \leq \theta_0\} + \beta 1\{0 \leq Z_i > \theta_0\} + \epsilon_i,$$

with the unknown parameters α, β, θ_0. What properties of such estimator can you establish (and if so, provide details of the proof).

Date: May 12, 2021.

QUAL EXAM: MATH 281C – SPRING 2021

- **Pick ONE** *from the following two questions to answer. If you choose to do both, the final score will be the maximum of the two (***NOT** *sum). Choose wisely before you start.*
- *Define any symbol you use unless its meaning is clear from context. Name any result you use if it has a name. Be concise and clear.* **Justify all your answers**.

3. Problem 3. (Hypothesis Testing for Bivariate Normal Family)

Assume $(X_1, Y_1), \dots, (X_n, Y_n)$ $(n \geq 2)$ are independent random variables from a bivariate normal distribution with pdf

$$\begin{aligned} f(x,y) &= f(x,y;\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho) \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{(x-\mu_1)^2}{2\sigma_1^2(1-\rho^2)} + \frac{\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)} - \frac{(y-\mu_2)^2}{2\sigma_2^2(1-\rho^2)} \right\}, \end{aligned}$$

where $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1^2, \sigma_2^2 > 0$ and $\rho \in (-1, 1)$.

(a) Find the conditional distribution of Y_i given $X_i = x$, that is, $Y_i | X_i = x$.

(b) Assume $\mu_1 = \mu_2 = 0$, and $\sigma_1^2, \sigma_2^2, \rho$ are unknown. Find the UMP unbiased test of level α for testing $H_0 : \rho = 0$ versus $H_1 : \rho \neq 0$. Describe the rejection region and the corresponding critical value.

3*. Problem 3*. (Hypothesis Testing for Negative Binomial)

Independent trials with constant probability p of success are carried out until a preassigned number m of successes has been obtained. If the number of trials required is $X + m$, then X has the negative binomial distribution $Nb(p, m)$:

$$\mathbb{P}(X = x) = \binom{m+x-1}{x} p^m (1-p)^x, \quad x = 0, 1, 2, \dots.$$

Let X, Y be independently distributed according to negative binomial distributions $Nb(p_1, m)$ and $Nb(p_2, n)$ respectively, and let $q_i = 1 - p_i$.

(a) Show that there exists a UMP unbiased test for testing $H_0 : \theta = q_2/q_1 \leq \theta_0$, and hence also for testing $H_0' : p_1 \leq p_2$ versus $H_1' : p_1 > p_2$.

(b) Determine the conditional distribution required for testing $H_0' : p_1 \leq p_2$ when $m = n = 1$.