# Math 281 – Qualifying Exam – Fall 2020

*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.*

**Problem 1.** Suppose that we observe data in pairs $(X, Y) \in \mathbb{R}^d \times \{\pm 1\}$, where the data come from a logistic model with $X \sim P_0$ and

$$p_\theta(y|x) = \frac{1}{1 + e^{-y \cdot x^\intercal \theta_0}}.$$

Define the log-loss $\ell_\theta(y|x) = \log(1 + e^{-y \cdot x^\intercal \theta})$. Let $\hat{\theta}_n$ minimize the empirical logistic loss

$$L_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ell_\theta(Y_i|X_i) = \frac{1}{n} \sum_{i=1}^n \log(1 + e^{-Y_i X_i^\intercal \theta})$$

from pairs $(X_i, Y_i)$ drawn from the logistic model with parameter $\theta_0$. Assume that the covaraites $X_i \in \mathbb{R}^d$ are i.i.d. and satisfy $\mathbb{E}(X_i X_i^\intercal) = \Sigma$ is positive definite and $\mathbb{E}\|X_i\|_2^4 < \infty$.

(a) Let $L(\theta) = \mathbb{E}_{\theta_0}\{\ell_\theta(Y|X)\}$ be the population logistic loss. Describe conditions under which $\theta_0$ is the unique minimizer of $\theta \mapsto L(\theta)$, that is, $\theta_0 \in \arg\min_{\theta \in \mathbb{R}^d} L(\theta)$.

(b) Under these assumptions show that $\hat{\theta}_n$ is consistent estimator of $\theta_0$ as $n \to \infty$. Provide details of your work. Hint: You may use the following fact about convex functions. For any convex function $h$, if there is some $r > 0$ and a point $x_0$ such that $h(x) > h(x_0)$ for all $x$ such that $\|x - x_0\|_2 = r$. Then $h(x') > h(x_0)$ for all $x'$ with $\|x' - x_0\|_2 > r$.

(c) Find the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n - \theta_0)$, provided that it is consistent. Describe the asymptotic covariance matrix of $\hat{\theta}_n$. Provide a heuristic argument to justify your findings.

**Problem 2.** Suppose that $X_1, \ldots, X_n$ is a random sample from a normal distribution with unit variance. Let $F(t) = \mathbb{P}(X_1 \leq t)$ be the cumulative distribution function (CDF) of $X_1$, and let $F_n(\cdot)$ be the empirical CDF. Given some $t \in \mathbb{R}$,

(a) Show that both $F_n(t)$ and $\Phi(t - \bar{X}_n)$ are consistent estimators of $F(t)$, where $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$;

(b) Compare the asymptotic variance of the estimators $F_n(t)$ and $\Phi(t - \bar{X}_n)$.

*Refer by number any result from the reference sheet that you use.*

**Problem 3.** Suppose we observe $X \sim \text{Bin}(n, \theta)$, where $n$ is known, and our goal is to estimate $\theta$. We use the *squared error loss*. (Note that an estimator $\delta$ is here simply a function from $\{0, \ldots, n\}$ to $\mathbb{R}$.)

(a) Derive an optimal estimator among unbiased estimators. Is it unique in that regard?

In what follows, we measure performance of an estimator $\delta$ based on the following average risk

$$r(\delta) = \int_0^1 \mathbb{E}_\theta[(\delta(X) - \theta)^2] w(\theta) \mathrm{d}\theta, \quad \text{where } w(\theta) = \frac{1}{\theta(1-\theta)}$$

(b) Is the prior proper or improper? Explain.

(c) Show that $r(\delta) < \infty$ if and only if $\delta(0) = 0$ and $\delta(n) = 1$.

(d) Derive an estimator that minimizes this average risk. Is it unique in that regard?