August 2021 Name:________________________

COMPREHENSIVE WRITTEN EXAM – STOR655 MATHEMATICAL STATISTICS

Unless otherwise noted, all problem parts have equal weight. In budgeting your time expect that some part will take longer than others. When solving multi-part problems feel free to use results of earlier parts even if you cannot solve them in proving later parts.

1. Let $X_1, X_2, \ldots$ be i.i.d. sample from a Pareto(λ, γ), i.e.,

$$f(x|\lambda,\gamma) = \frac{\gamma\lambda}{(1+\lambda x)^{\gamma+1}} I_{(0,\infty)}(x), \quad \lambda > 0, \gamma > 0.$$

In some of the problem parts below you might need to make extra assumptions? If you make extra assumptions, state them clearly.

(a) Does this two parameter Pareto(λ, γ) form an exponential family?

(b) Find the Method of Moments (MM) estimator $\hat{\theta}_{MM}$ based on first two moments.

(c) Is $\hat{\theta}_{MM}$ strongly consistent?

(d) Is $\hat{\theta}_{MM}$ asymptotically normal? If yes, what is its asymptotic variance?

(e) Is $\hat{\theta}_{MM}$ asymptotically efficient? If it is not, how would you improve it by scoring?

For the rest of this problem we assume $\lambda = 1, \gamma > 2$ and consider the U-statistics:

$$U_n = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \left(X_i(1+X_j)\log(1+X_j) + X_j(1+X_i)\log(1+X_i) \right)$$

(f) Find $\eta = EU_n$.

(g) Find Hájek projection $\pi_n(U_n - \eta)$.

(h) Does $\sqrt{n}(U_n - \eta) \xrightarrow{\mathcal{D}} N(0, \xi^2)$? If yes, find ξ^2.

Hint: The following integrals might be helpful throughout:

For $\gamma > 1$

$$\int_0^\infty \frac{x\,\gamma}{(1+x)^{\gamma+1}}\,dx = \frac{1}{\gamma-1}, \qquad \int_0^\infty \frac{(1+x)\log(1+x)\gamma}{(1+x)^{\gamma+1}}\,dx = \frac{\gamma}{(\gamma-1)^2}$$

For $\gamma > 2$

$$\int_0^\infty \frac{x^2\,\gamma}{(1+x)^{\gamma+1}}\,dx = \frac{2}{(\gamma-1)(\gamma-2)}$$

2. Let $X_1, \ldots, X_n$ be i.i.d. sample from Negative Binomial$(2, p)$ distribution, i.e.

$$f(x|p) = (x+1)p^2(1-p)^x I_{\{0,1,\ldots\}}(x).$$

Consider the $N(0,1)$ prior on the log odds ratio $\psi = \log(p/(1-p))$. What is a suitable Gaussian approximation for the posterior $p(\psi|x_1, \ldots, x_n)$ for large n? In what sense is this approximation asymptotically valid?