Question I. $(X, \mathcal{F})$ is a statistical model.

data

(a) What is meant by saying $Y = Y(x)$ is a "complete" data reduction?

Usually in our course, when x was iid. real data from a nonparametric family, the order statistics were not only sufficient but complete too. This is not always so however; the argument we used does not extend to all nonparametric families.

(b) Explain which of the following families of parent distributions on $\mathbb{R}$ lead to completeness of the order statistics $(X_{(1)}, \dots, X_{(n)})$ in iid. sampling. It is possible that your answers may depend on the sample size n.

(i) $\{$continuous symmetric distributions with known center of symmetry $\theta_0\}$.

(ii) $\{$continuous distributions with known unique median $\theta_0\}$.

(iii) $\{$continuous distributions with a center of symmetry$\}$

Question II. $U_1, \dots, U_n \sim$ iid unif $[\theta-1, \theta]$; $\theta \in \mathbb{R}$; $n \geq 2$

(a) Explain whether this is a location family; what about a scale family.

(b) Show that $(U_{(1)}, U_{(n)})$ is sufficient but not complete.

(c) Show that $(U_{(1)} + U_{(n)})/2$ is not sufficient.

(d) Show that $\bar{u} + \frac{1}{2}$ is not UMVUE for θ.

Question III

(a) Define what is meant by a least-favorable sequence of priors (Λ_m).

(b) Prove the following (familiar) theorem. (The loss function is fixed.)

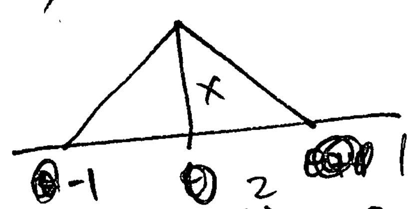
"Let (Λ_m) be a sequence of priors with corresponding Bayes risks of their Bayes rules $r_m \to r$. Let δ be an estimator with $\sup_\theta R_\theta(\delta) = r$; then δ is minimax."
[Also (Λ_m) is least-favorable, but no need to prove that.]

We used this theorem to prove minimaxity of $\bar{X}$ when $X_1, \dots, X_n$ iid $\mathcal{N}(\mu, \sigma^2)$ with squared-error loss. The proof involved an artificiality in that we needed to restrict the parameter space by bounding σ^2.

(c) Why?

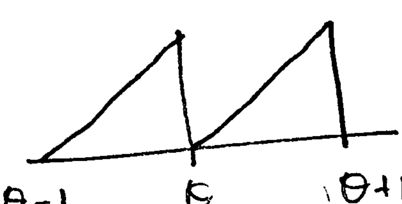
(d) Show with a modified loss function $L(d, \mu) = \frac{(d-\mu)^2}{\sigma^2}$, the need for that artificiality goes away, and show, using a <u>small</u> modification of the argument we saw, that $\bar{x}$ is once again minimax.

Question IV $X_1, \dots, X_n$ iid f_θ :



(a) Find the AR(LL)E of the mean to the median for this problem

Now change f_θ :



(b) Show that the median is no longer $\sqrt{n}$-consistent.

(c) What is its consistency rate?

(d) What is its weak limit (limiting distribution) (with suitable re-scaling)?

(e) Which is the asymptotically preferable estimator now: the sample mean or the sample median?

Question V. Consider a regular one-parameter likelihood problem for estimating $\theta \in \mathbb{R}^1$.

(a) State the theorem on efficient likelihood estimation (ELE). (No need to spell out the regularity conditions precisely.)

(b) Explain briefly and without technicality how it can happen that the sequence of ELE's may not be observable (i.e. statistics).

(c) Prove that if there is a consistent estimator δ_n for the target θ, and you pick the root $\hat{\theta}_n$ (from the root set of the likelihood equation) closest to δ_n, then $(\hat{\theta}_n)$ is an ELE which is observable.

Problem 1. Let $X_1, \dots, X_n$ be i.i.d. from the gamma distribution $\Gamma(g, b)$, which has density

$$f(x) = C(g,b)x^{g-1}e^{-x/b}, \quad x > 0,$$

where $C(g, b)$ is a normalizing constant. Note that $g, b > 0$.

Find a UMP level α test (if it exists) in the following two situations:

10 (a) $H : b \le b_0$ versus $K : b > b_0$, and g is known.

10 (b) $H : g \le g_0$ versus $K : g > g_0$, and b is known.

Be as specific as you can.

Problem 2. Consider the following distributions on $\{1, \dots, 5\}$:

	1	2	3	4	5
P_1	0.3	0.2	0.3	0.1	0.1
P_2	0.2	0.2	0.1	0.4	0.1
Q	0.2	0.0	0.3	0.2	0.3

We want to test $\{P_1, P_2\}$ versus Q.

5 1. Find an MP test at level $\alpha = 0.05$. Is it unique?

10 2. Find an MP test at level $\alpha = 0.15$. Is it unique?

10 3. Find an MP test at level $\alpha = 0.50$. Is it unique?

5 4. Is there a level α for which the MP test not unique? Justify your answer.

Below is a plot of

$$L_\lambda(x) = \frac{Q(x)}{(1-\lambda)P_1(x) + \lambda P_2(x)}$$

Each line corresponds to $L_\lambda(x)$ as λ varies from 0 to 1, for each $x = 1, 2, 3, 4, 5$.