Question I. $U_1, \ldots, U_n$ – iid unif$[\beta, \beta+1]$; $\beta \in \mathbb{R}$; $n \geq 2$

(a) Explain whether this is a location family, or a scale family, or perhaps both.

(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 UMVU for β

(e) Find the ARE of $\bar{u} - \frac{1}{2}$ ($= \hat{\beta}_1$ say) to median$\{X_i\} - \frac{1}{2}$ ($= \hat{\beta}_2$ say) as estimators of β in this problem.

(f) Derive the MLE of β for this problem. What is its consistency rate? Support your answer with a rough argument, and if the rate is not the usual $\sqrt{n}$-rate, what regularity conditions are violated to make the usual theorem's conclusion invalid?

Question II. A toy testing problem: $\Theta = \{0, 1\}$, Z is data taking values in a 3-point set $\{1, 2, 3\}$.

Null H_0 : $\theta = 0$; $Z \sim f_0(z)$
Alternative H_1 : $\theta = 1$; $Z \sim f_1(z)$

discrete distributions, tabulated below.

	$z = 1$	$z = 2$	$z = 3$
$f_0(z)$	$\frac{1}{4}$	0	$\frac{3}{4}$
$f_1(z)$	$\frac{1}{2}$	$\frac{1}{2}$	0

(a) Find the most powerful test of H_0 vs. H_1 at level $\alpha = .25$, and find its power.

(b) Why does a most powerful test (simple vs. simple) always have power at least as large as its significance level?

(c) Find a test based on the likelihood ratio (as in (a)) with level $\alpha = .3$. It will be a randomized test. Explain how you perform it using Z and a coin.

(d) Adjoin another distribution f_2 to the alternative hypothesis (still supported in $\{1, 2, 3\}$) so your test in (a) is still UMP against $H_2 = \{f_1, f_2\}$.

(e) Adjoin a third distribution f_3 so your test in (a) is no longer UMP against $\{f_1, f_2, f_3\}$.

(f) A Bayesian sets a prior π on Θ with $\pi(0) = .4$; $\pi(1) = .6$

Derive the Bayes rule for estimating θ with squared-error loss, and compute the Bayes risk of the Bayes rule.

(g) Is the prior in (f) least-favorable? What is?

Question III. (On skewness estimation)

Recall that when $X_1, \ldots, X_n \sim$ iid (μ, σ^2) $(E|X_1|^3 < \infty)$ the skewness is defined by $\tau = E[X_1 - \mu]^3 / \sigma^3$. ($\tau$ is a location- and scale-invariant measure of asymmetry, and its estimation is useful in fitting "Edgeworth corrections" to the CLT approximation.)

The usual estimator under nonparametric assumptions is

$$\hat{\tau} = \hat{\tau}_n = \frac{\frac{1}{n}\Sigma(X_i - \bar{X})^3}{\left(\frac{1}{n}\Sigma(X_i - \bar{X})^2\right)^{3/2}} = \frac{\hat{\mu}_3}{\hat{\sigma}_3} \text{ say.}$$

(a) How do we know this is not UMVUE under normality assumptions?

(b) Give an example of a discrete bivariate distribution, supported on the smallest set you can think of, where the one coordinate r.v. is either independent of, or uncorrelated with, the other, but not both.

(c) Show that under normality the numerator $\hat{\mu}_3$ and denominator $\hat{\sigma}^3$ in $\hat{\tau}$ above have the property in (b) (although of course this is not a discrete situation).

(d) How do we know readily that under normal sampling $\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{d} N(0, 2\sigma^4)$? (There are several possible answers.)

(e) Find a similar weak convergence result for $\hat{\mu}_3$. (The binomial theorem and Slutsky's theorem may come in handy. Also, you may assume that $EZ^6 = 15$ when $Z \sim \mathcal{N}(0,1)$.)

(f) In class we mentioned without proof a bivariate form of the delta method. Roughly, if $g(\theta_1, \theta_2)$ is nice, and

$$\sqrt{n}\left(\begin{pmatrix}\hat{\theta}_1\\ \hat{\theta}_2\end{pmatrix} - \begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}\right) \xrightarrow{d} \mathcal{N}_2\left(\underline{0}, \begin{pmatrix}\sigma_{11}^2 & \sigma_{12}\\ \sigma_{12} & \sigma_{22}^2\end{pmatrix}\right),$$ then

$$\sqrt{n}\left(g(\hat{\theta}_1, \hat{\theta}_2) - g(\theta_1, \theta_2)\right) \xrightarrow{d} \mathcal{N}\left(0,\ \sigma_1^2\left(\frac{\partial g}{\partial \theta_1}(\theta_1, \theta_2)\right)^2 + \sigma_2^2\left(\frac{\partial g}{\partial \theta_2}(\theta_1, \theta_2)\right)^2 + 2\sigma_{12}\frac{\partial g}{\partial \theta_1}(\theta_1, \theta_2)\cdot\frac{\partial g}{\partial \theta_2}(\theta_1, \theta_2)\right)$$

Using this formula, and simplifying it by assuming $\sigma_{12} = 0$ (because of what you found in (c)), show that, if in fact your original data were from $\mathcal{N}(\mu, \sigma^2)$, $\sqrt{n}\,\hat{\tau}_n \xrightarrow{d} \mathcal{N}(0, 6\sigma^4)$.