Please provide complete and well-written solutions to the following exercises.

Due October 10, 6PM, to be submitted in brightspace, under the Content tab (so it actually is submitted to gradescope).

Homework 3

Exercise 1. Let $X_1, \ldots, X_n$ be i.i.d. Gaussian random variables with unknown mean and unknown variance.

- Find a real-valued pivotal quantity for $X = (X_1, \ldots, X_n)$.
- Using the pivotal quantity, construct a $1 - \alpha$ confidence interval for the mean μ, for any $0 < \alpha < 1$.

Exercise 2. Let $X_1, \ldots, X_n$ be a real-valued random sample of size n from a family of distributions $\{f_\theta \colon \theta \in \Theta\}$. Suppose $\Theta = \mathbf{R}$. Fix $\theta \in \mathbf{R}$. Denote $X := (X_1, \ldots, X_n)$. Consider a set of nonrandomized hypothesis tests with rejection regions $C_\alpha \subseteq \mathbf{R}^n$ for all $\alpha \in [0, 1]$. Suppose these rejection regions are nested in the sense that $C_\alpha \subseteq C_{\alpha'}$ for all $0 \leq \alpha < \alpha' \leq 1$. As usual, denote $\Theta = \Theta_0 \cup \Theta_1$ with $\Theta_0 \cap \Theta_1 = \emptyset$. Define also the p-valued $p(x) := \inf\{\alpha \in [0, 1] \colon x \in C_\alpha\}$, $\forall\, x \in \mathbf{R}^n$.

- Suppose $\sup_{\theta \in \Theta_0} \mathbf{P}_\theta(X \in C_\alpha) \leq \alpha$ for all $0 \leq \alpha \leq 1$. Show that the p-valued satisfies
$$\mathbf{P}_\theta(p(X) \leq c) \leq c, \qquad \forall\, 0 \leq c \leq 1,\ \forall\, \theta \in \Theta_0.$$
- Suppose $\mathbf{P}_\theta(X \in C_\alpha) = \alpha$ for all $0 \leq \alpha \leq 1$. Show that the p-valued satisfies
$$\mathbf{P}_\theta(p(X) \leq c) = c, \qquad \forall\, 0 \leq c \leq 1,\ \forall\, \theta \in \Theta_0.$$
That is, $p(X)$ is uniformly distributed in $[0, 1]$.

Exercise 3. Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution with unknown location parameter $\theta > 0$, i.e. X_1 has density
$$g(x) := 1_{x \geq \theta} e^{-(x - \theta)}, \qquad \forall\, x \in \mathbf{R}.$$

Fix $\theta_0 \in \mathbf{R}$. Suppose we want to test that hypothesis H_0 that $\theta \leq \theta_0$ versus the alternative H_1 that $\theta > \theta_0$. That is, $\Theta = \mathbf{R}$, $\Theta_0 = \{\theta \in \mathbf{R} \colon \theta \leq \theta_0\}$ and $\Theta_0^c = \Theta_1 = \{\theta \in \mathbf{R} \colon \theta > \theta_0$.

- Explicitly describe the rejection region of the generalized likelihood ratio test for this hypothesis. (Hint: it might be easier to describe the region using $x_{(1)} = \min(x_1, \ldots, x_n)$.)
- Prove that $X_{(1)} := \min(X_1, \ldots, X_n)$ is a sufficient statistic for θ.
- (Optional) If H_0 is true, then does
$$2 \log \frac{\sup_{\theta \in \Theta} f_\theta(X_1, \ldots, X_n)}{\sup_{\theta \in \Theta_0} f_\theta(X_1, \ldots, X_n)}$$

converge in distribution to a chi-squared distribution as $n \to \infty$?

Exercise 4. Let $X_1, \ldots, X_n$ be a random sample from a Gaussian random variable with unknown mean $\mu \in \mathbf{R}$ and unknown variance $\sigma^2 > 0$.

Fix $\mu_0 \in \mathbf{R}$. Suppose we want to test that hypothesis H_0 that $\mu = \mu_0$ versus the alternative H_1 that $\mu \neq \mu_0$.

- Explicitly describe the rejection region of the generalized likelihood ratio test for this hypothesis.
- Give an explicit formula for the p-value of this hypothesis test. (Hint: If S^2 denotes the sample variance and $\overline{X}$ denotes the sample mean, you should then be able to use the statistic $\frac{(\overline{X}-\mu_0)^2}{S^2}$. Since we have an explicit formula for Snedecor's distribution, you should then be able to write an explicit integral formula for the p-value of this test.)