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

Due September 1, 9AM, to be submitted in blackboard, under the Assignments tab.

Homework 1

Exercise 1. Estimate the probability that 1000000 coin flips of fair coins will result in more than $501,000$ heads, using the Central Limit Theorem. (Some of the following integrals may be relevant: $\int_{-\infty}^{0} e^{-t^2/2}dt/\sqrt{2\pi} = 1/2$, $\int_{-\infty}^{1} e^{-t^2/2}dt/\sqrt{2\pi} \approx .8413$, $\int_{-\infty}^{2} e^{-t^2/2}dt/\sqrt{2\pi} \approx .9772$, $\int_{-\infty}^{3} e^{-t^2/2}dt/\sqrt{2\pi} \approx .9987$.) (Hint: use Bernoulli random variables.)

Exercise 2 (Numerical Integration)**.** In computer graphics in video games, etc., various integrations are performed in order to simulate lighting effects. Here is a way to use random sampling to integrate a function in order to quickly and accurately render lighting effects. Let $\Omega = [0, 1]$, and let $\mathbf{P}$ be the uniform probably law on Ω, so that if $0 \leq a < b \leq 1$, we have $\mathbf{P}([a, b]) = b - a$. Let $X_1, \ldots, X_n$ be independent random variables such that $\mathbf{P}(X_i \in [a, b]) = b - a$ for all $0 \leq a < b \leq 1$, for all $i \in \{1, \ldots, n\}$. Let $f\colon [0, 1] \to \mathbf{R}$ be a continuous function we would like to integrate. Instead of integrating f directly, we instead compute the quantity

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i).$$

Show that

$$\lim_{n\to\infty} \mathbf{E}\left(\frac{1}{n}\sum_{i=1}^{n} f(X_i)\right) = \int_0^1 f(t)dt.$$

$$\lim_{n\to\infty} \mathrm{var}\left(\frac{1}{n}\sum_{i=1}^{n} f(X_i)\right) = 0.$$

That is, as n becomes large, $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$ is a good estimate for $\int_0^1 f(t)dt$.

Exercise 3. Let $X := (X_1, \ldots, X_n)$ be a random sample of size n from a binomial distribution with parameters n and p. Here n is a positive (known) integer and $0 < p < 1$ is unknown. (That is, $X_1, \ldots, X_n$ are i.i.d. and X_1 is a binomial random variable with parameters n and p, so that $\mathbf{P}(X_1 = k) = \binom{n}{k}p^k(1-p)^{n-k}$ for all integers $0 \leq k \leq n$.)

You can freely use that $\mathbf{E}X_1 = np$ and $\mathrm{Var}X_1 = np(1-p)$.

- Computer the Fisher information $I_X(p)$ for any $0 < p < 1$.
  (Consider n to be fixed.)
- Let Z be an unbiased estimator of p^2 (assume that Z is a function of $X_1, \ldots, X_n$). State the Cramér-Rao inequality for Z.
- Let W be an unbiased estimator of $1/p$ (assume that W is a function of $X_1, \ldots, X_n$). State the Cramér-Rao inequality for W.

Exercise 4. Let $X_1, \ldots, X_n$ be a random sample of size n from a Poisson distribution with unknown parameter $\lambda > 0$. (So, $\mathbf{P}(X_1 = k) = e^{-\lambda}\lambda^k/k!$ for all integers $k \geq 0$.)

- Find an MLE (maximum likelihood estimator) for λ.
- Is the MLE you found unique? That is, could there be more than one MLE for this problem?

Exercise 5. Suppose X is a Gaussian distributed random variable with known variance $\sigma^2 > 0$ but unknown mean. Fix $\mu_0, \mu_1 \in \mathbf{R}$. Assume that $\mu_0 - \mu_1 > 0$. We want to test the hypothesis H_0 that $\mu = \mu_0$ versus the hypothesis H_1 that $\mu = \mu_1$. Fix $\alpha \in (0, 1)$. Explicitly describe the UMP test for the class of tests whose significance level is at most α.

Your description of the test should use the function $\Phi(t) := \int_{-\infty}^{t} e^{-x^2/2}dx/\sqrt{2\pi}$, $\Phi \colon \mathbf{R} \to (0, 1)$, and/or the function $\Phi^{-1} \colon (0, 1) \to \mathbf{R}$. (Recall that $\Phi(\Phi^{-1}(s)) = s$ for all $s \in (0, 1)$ and $\Phi^{-1}(\Phi(t)) = t$ for all $t \in \mathbf{R}$.)

Exercise 6. This exercise demonstrates that a UMP might not always exists.

Let $X_1, \ldots, X_n$ be i.i.d. Gaussian random variables with known variance and unknown mean $\mu \in \mathbf{R}$. Fix $\mu_0 \in \mathbf{R}$. Let H_0 denote the hypothesis $\{\mu = \mu_0\}$ and let H_1 denote the hypothesis $\mu \neq \mu_0$. Fix $0 < \alpha < 1$. Let $\mathcal{T}$ denote the set of hypothesis tests with significance level at most α. Show that no UMP class $\mathcal{T}$ test exists, using the following strategy.

- Let $\mu_1 < \mu_0$. You may take as given the following fact (that follows from the Karlin-Rubin Theorem): the power at μ_1 is maximized among class $\mathcal{T}$ tests by the hypothesis test ϕ that rejects H_0 when the sample mean satisfies $\overline{X} < c$ for an appropriate choice of $c \in \mathbf{R}$. Assume for the sake of contradiction that a UMP class $\mathcal{T}$ test ϕ' exists. Then, using the necessity part of the Neyman-Pearson Lemma (i.e. consider testing $\mu = \mu_0$ versus $\mu = \mu_1$), conclude that ϕ' must have the same rejection region as ϕ (just by examining the power of the tests at μ_1.)
- Consider now a test in $\mathcal{T}$ that rejects H_0 when the sample mean satisfies $\overline{X} > c'$ for an appropriate choice of $c' \in \mathbf{R}$. Repeating the previous argument, conclude that ϕ' must reject when $\overline{X} > c'$, leading to a contradiction.

  That is, let $\mu_2 > \mu_0$. You may take as given the following fact (that follows from the Karlin-Rubin Theorem): the power at μ_2 is maximized among class $\mathcal{T}$ tests by the hypothesis test ϕ'' that rejects H_0 when the sample mean satisfies $\overline{X} > c'$ for an appropriate choice of $c' \in \mathbf{R}$. Then, using the necessity part of the Neyman-Pearson Lemma (i.e. consider testing $\mu = \mu_0$ versus $\mu = \mu_2$), conclude that ϕ' must have the same rejection region as ϕ''.

Exercise 7. The rejection regions C_α for UMP hypothesis tests of significance level at most $\alpha \in (0, 1)$ are often nested in the sense that $C_\alpha \subseteq C_{\alpha'}$ for all $0 < \alpha < \alpha' < 1$. This exercise demonstrates an example of UMP tests where this nesting behavior does not occur.

Let $\theta_0, \theta_1 \in \mathbf{R}$ be unequal parameters. Let H_0 denote the hypothesis $\{\theta = \theta_0\}$ and let H_1 denote the hypothesis $\{\theta = \theta_1\}$. Suppose $X \in \{1, 2, 3\}$ is a random variable. If $\theta = \theta_0$, assume that X takes the values $1, 2, 3$ with probabilities $.85, .1, .05$, respectively. If $\theta = \theta_1$, assume that X takes the values $1, 2, 3$ with probabilities $.7, .2, .1$, respectively. Let $\mathcal{T}$ denote the set of hypothesis tests with significance level at most α.

- Let $0 < \alpha < .15$. Show that a UMP class $\mathcal{T}$ test is not unique.
- When $\alpha = .05$, show there is a unique nonrandomized hypothesis UMP class $\mathcal{T}$ test.
- When $\alpha = .1$, show there is a unique nonrandomized hypothesis UMP class $\mathcal{T}$ test.
- Show that the $\alpha = .05$ and $\alpha' = .1$ UMP nonrandomized tests from above do not have nested rejection regions.
- However, when $\alpha = .05$ and $\alpha' = .1$, there are randomized UMP tests $\phi, \phi' \colon \mathbf{R}^n \to [0, 1]$ respectively, that are nested in the sense that $\phi \leq \phi'$.