Spring 2026: Mathematical Statistics
Recitation Worksheet 9
April 3, 2026

Midterm 2 Short Review Problems

1. True or False?
   (a) In order to find the maximum likelihood estimator, it suffices to find the estimator that achieves zero gradient/derivative of the likelihood function.
   (b) The MLE estimator $\hat{\theta}$ has distribution $N(\theta, \frac{1}{nI(\theta)})$, where $I(\theta)$ is the Fisher information.
   (c) A biased estimator can achieve lower than $\frac{1}{nI(\theta)}$ variance.
   (d) The method of moments estimator is consistent if the moments are a continuously invertible function of θ.
   (e) In the Bayesian setting, the posterior mode converges to the MLE in the limit as the number of data points $n \to \infty$ for any choice of prior.
   (f) There is a tradeoff between the power of a test and its significance level.
   (g) If a test rejects the null hypothesis with probability 0.5 regardless of the data, then the significance level is 0.5.
   (h) For a one-sided Generalized LR test, e.g. $H_0 : \mu = \mu_0$, $H_1 : \mu > \mu_0$, we reject if $LR < c$. For a two-sided Generalized LR test, e.g. $H_1 : \mu \neq \mu_0$, we reject if $LR < -c$ or $LR > c$.
   (i) As the number of data points $n \to \infty$, we always expect to see $-2 \log LR$ converge in distribution to a chi-squared distribution.
2. Short Answer
   (a) Suppose we measure the heights of 10 students in this class. What is an unbiased estimate of the population variance if the students are sampled with replacement? What if they're sampled without replacement?
   (b) Suppose we want a CI for the average height. (You may assume that height is a normally distributed variable, with unknown mean and variance). What CDF table would we look in?
3. Suppose we want to test if a disease has any effect on sleep. We collect data on N independently selected people, recording whether they have the disease, and whether their average length of sleep is < 5 hours, $5 - 7$ hours, $7 - 9$ hours, or > 9 hours. Our null hypothesis is that the probability of sleep length falling into each category is the same whether one has the disease or not, and the alternative hypothesis is that the probability is not the same. If we use the generalized LR test with the fact that $-2 \log LR \to \chi^2_d$, what is d?
4. Suppose the data set $X_1, X_2, \ldots, X_n$ is drawn iid from Unif($[0, N]$). What is the likelihood function? Find the MLE of N.

From Practice Midterms

1. [2025 Problem 1(c)] True or False? Let $X_1, \ldots, X_n \sim \text{Multi}(n, \bar{p})$ follow a multinomial distribution with probability vector $\bar{p} = (p_1, \ldots, p_m)$. Let $H_0 : \bar{p} \in \text{span}\{\bar{p}_0, \bar{p}_0'\}$ and $H_1 : \bar{p} \notin \text{span}\{\bar{p}_0, \bar{p}_0'\}$. Then $-2 \log LR \xrightarrow{d} \chi^2_{m-1}$ as $n \to \infty$.

2. [2025 Problem 5] Carol studies whether the "Big Three" (Djokovic, Federer, and Nadal) are equally dominant in their rivalries. She collects the following match data:

Rivalry	Number of matches	Number of wins
Djokovic–Federer	50	27
Djokovic–Nadal	60	31
Federer–Nadal	40	24

   Assume each row follows an independent binomial distribution: $X_{D-F} \sim B(n = 50, p_{D-F})$, $X_{D-N} \sim B(n = 60, p_{D-N})$, and $X_{F-N} \sim B(n = 40, p_{F-N})$, with unknown probabilities $(p_{D-F}, p_{D-N}, p_{F-N})$. Carol tests:

$$H_0 : p_{D-F} = p_{D-N} = p_{F-N} = 0.5 \text{ v.s. } H_1 : \text{not } H_0.$$

   Carry out the generalized likelihood ratio test, by writing down the expression for the test statistic $-2 \log LR$, and identifying its approximate null distribution. You may use without proof that the MLE of p for a binomial observation $X \sim B(n, p)$ is $\hat{p} = X/n$. No numerical computation is needed.

3. [2025 Problem 3] Based on n i.i.d. observations $X_1, \ldots, X_n$, and under an assumed statistical model, Alice computes that the log-likelihood function $\ell_n(\theta)$ is approximately

$$\frac{\ell_n(\theta)}{n} \approx 1 - 2(\theta - 3)^2.$$

   Based on this expression, write down the approximate value of the MLE $\hat{\theta}_n$, and the approximate distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$. (Hint: how is $\ell_n''(\theta)/n$ related to the Fisher information $I(\theta)$?)