APMA 1650

Homework 9

Instructions: Homework is due by 11:59pm EST in Gradescope on the day listed on the course webpage. You can use calculators for this assignment. Solutions must be written independently and *cannot* be shared with any other students.

You must *show all work* and *explain your answers thoroughly* to get full credit. **You will be graded partly on how well you explain the answers**.

There will be a 5pt penalty for homework submitted with problems incorrectly assigned to a page. A 15pt penalty will be applied for homework submitted during the late window.

Problem 1. (16 pts) Suppose you are trying to measure the temperature θ of a pot of water down to the nearest microKelvin (you don't need to know what microKelvin are). The thermometer you have tends to produce random errors in the measurements down to that scale. To combat this, you make n measurements of the temperature $X_1, X_2, \ldots X_n$. Suppose you know that each measurement has a random error E_i given by

$$X_i = \theta + E_i$$

and that $E_1, E_2, \ldots E_n$ are iid with $E[E_i] = 0$ and $\text{Var}(E_i) = 9$. You decide to average your measurements to get a better idea of the true temperature,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

a. (8 pts) Use Chebyshev to estimate how large n needs to be at least 90% sure that the error $|\overline{X} - \theta|$ is with in 0.2 units.

b. (8 pts) Improve your answer to (b) instead using the CLT (round your answer to the nearest integer).

Problem 2. (24 pts) Let $X_1, X_2, \ldots X_{100}$ be iid standard normal random variables. Let

$$Y = X_1^2 + X_2^2 + \ldots + X_{100}^2.$$

a. (8 pts) Show that X_i^2 is $\text{Gamma}(\alpha, \beta)$ distributed for certain α and β (you may use the fact that $\Gamma(1/2) = \sqrt{\pi}$).

b. (5 pts) Use the known mean and variance of the Gamma distribution and part (a) to determine EX_i^2 and $\text{Var}(X_i^2)$.

c. (8 pts) Use the Central Limit Theorem to find a value y such that $P(Y > y) \approx 0.05$. Approximate to 3 decimal places.

d. (5 pts) It is well known that Y has a $\chi^2(n)$ distribution, compare your answer in part (c) to the one obtained via the $\chi^2(n)$ table provided on the course webpage (or using `chi2inv` in MATLAB)

Problem 3. (30 pts) Let $X_1, X_2, \ldots X_n$ be a random sample from a distribution associated to random variable X with parameter θ.

a. (10 pts) If the distribution of X has PDF

$$f_X(x;\theta) = \begin{cases} (\theta+1)x^{\theta} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

For $\theta > -1$. Find the maximum likelihood estimator $\hat{\Theta}_{MLE}$ of θ.

b. (10 pts) If the distribution of X is Poisson, find the maximum likelihood estimator $\hat{\Theta}_{MLE}$ for $\theta = E[2^X]$.

c. (8 pts) Use the WLLN and properties of convergence in probability to show that $\hat{\Theta}_{MLE}$ is a consistent estimator of θ for both (a) and (b) above.

Problem 4. (40 pts) Let $X_1, X_2, \ldots, X_n$ be a random sample from Uniform$(0, \theta)$.

a. (8 pts) Show that the order statistic

$$\hat{\Theta} = \max\{X_1, X_2, \ldots, X_n\}$$

is a maximum likelihood estimator for θ. (Hint: taking derivatives isn't always the best strategy to finding a maximum)

b. Show that $\hat{\Theta}$ in part (a) has a CDF given by

$$F_{\hat{\Theta}}(x) = \begin{cases} 0 & x \le 0 \\ \left(\frac{x}{\theta}\right)^n & x \in [0, \theta] \\ 1 & \text{otherwise} \end{cases}$$

(Hint: $\{\hat{\Theta} \le x\} \Leftrightarrow \{X_1 \le x\} \cap \{X_2 \le x\} \cap \ldots \cap \{X_n \le x\}$)

c. (6 pts) What is the bias $B(\hat{\Theta})$? Explain intuitively why this makes sense.

d. (8 pts) Find a value c so that $\tilde{\Theta} = c\hat{\Theta}$ is unbiased.

e. (8 pts) Find the mean square errors $MSE(\hat{\Theta})$, $MSE(\tilde{\Theta})$. Which one is better?

Problem 5. (20 pts) Suppose you have two coins, one is fair, the other produces heads with probability 3/4. You are handed one of the coins and decide to flip the coin n times to try to figure out if it's fair or not.

a. (4 pts) Assuming you know which coin was chosen, explain what the WLLN predicts the proportion of heads should look like as you take n large.

b. (8 pts) Use Chebyshev to estimate the number of coin flips you need to take to be 95% sure you know which coin was chosen.

c. (8 pts) Use the CLT to estimate how many coin flips you need to take to be 95% sure you know which coin was chosen.

Problem 6. (20 pts) You are trying to guess the mean score on the latest midterm in a large class. You randomly sampled 10 students and ask them what their score was. The responses you get are

$$73, 82, 91, 50, 68, 77, 92, 81, 75, 69.$$

Assume the grades are normally distributed (just roll with it...) with unknown mean μ and variance σ^2.

a. (10 pts) Find a 90% confidence interval for the variance σ^2

b. (10 pts) Find a 95% confidence interval for the mean μ.

You may use one of the tables provided on the course webpage or `chi2inv` and `tinv` in MATLAB.