STA261 S19: Test 2 Solutions

No aids. 60 minutes. Write all answers directly beneath where the question is asked. Use pages 4 and 5 for rough work.

1. *Basic, 4 marks* Let $X_1, \ldots, X_n$ be an IID sample from a parametric family of distributions $\{F_\theta : \theta \in \Theta\}$ with corresponding densities f_θ.

a) (2) Define the Likelihood function in words and very briefly explain how it is used for inference on θ.

Solution. The Likelihood function evaluated at θ is the probability under our model of observing the data given that the true parameter is θ (1 mark). We can use it to get a region of θ values that correspond to a high probability under our model of observing the data that we did (1 mark).

b) (2) Give a mathematical definition of the following in terms of f_θ (.5 each):

Likelihood function:

Solution.

$$L(\theta \mid X_1, \ldots, X_n) = \prod_{i=1}^{n} f_\theta(X_i) \qquad \text{(0.5 marks)}$$

Score function:

Solution.

$$S(\theta \mid X_1, \ldots, X_n) = \frac{\partial \log L(\theta \mid X_1, \ldots, X_n)}{\partial \theta} = \frac{\partial \sum_{i=1}^{n} \log f_\theta(X_i)}{\partial \theta} \qquad \text{(0.5 marks)}$$

Observed Information:

Solution.

$$\widehat{I}(X_1, \ldots, X_n) = -\frac{\partial^2 \log L(\theta \mid X_1, \ldots, X_n)}{\partial \theta^2}\bigg|_{\theta=\theta_{MLE}} = -\frac{\partial^2 \sum_{i=1}^{n} \log f_\theta(X_i)}{\partial \theta^2}\bigg|_{\theta=\theta_{MLE}} \qquad \text{(0.5 marks)}$$

Expected/Fisher Information:

Solution.

$$I(\theta) = \mathrm{Var}_\theta\left(S(\theta \mid X_1, \ldots, X_n)\right) = \mathrm{Var}_\theta\left(\frac{\partial \sum_{i=1}^{n} \log f_\theta(X_i)}{\partial \theta}\right) \qquad \text{(0.5 marks)}$$

Also acceptable: negated expected second derivative of log likelihood.

2. *Adept, 4 marks.* Suppose $X_1, \dots, X_n$ is an IID sample from a Poisson(λ) distribution, having $P(X = x) = \lambda^x e^{-\lambda}/x!$ for $x = 0, 1, 2, \dots$

a) (1) Find the Likelihood function for λ.

Solution.

$$\begin{aligned} L(\lambda \mid X_1, \dots, X_n) &= \prod_{i=1}^{n} P_\lambda(X_i) \\ &= \prod_{i=1}^{n} \frac{\lambda^{X_i} e^{-\lambda}}{X_i!} \\ &= \frac{\lambda^{\sum_{i=1}^n X_i} e^{-n\lambda}}{\prod_{i=1}^n X_i!}. \end{aligned} \qquad \text{(1 mark)}$$

b) (1) Show that $T = \sum_{i=1}^n X_i$ is a sufficient statistic for λ. Briefly explain why this means that the sample mean $\bar{X} = T/n$ is also sufficient.

Solution. Consider two different IID samples $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ from the same Poisson model that satisfy $T = \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i$ (0.25 marks). Then, using 2.a),

$$\begin{aligned} \frac{L(\lambda \mid X_1, \dots, X_n)}{L(\lambda \mid Y_1, \dots, Y_n)} &= \frac{\lambda^{\sum_{i=1}^n X_i} e^{-n\lambda}}{\prod_{i=1}^n X_i!} \frac{\prod_{i=1}^n Y_i!}{\lambda^{\sum_{i=1}^n Y_i} e^{-n\lambda}} \\ &= \frac{\lambda^T e^{-n\lambda}}{\lambda^T e^{-n\lambda}} \frac{\prod_{i=1}^n Y_i!}{\prod_{i=1}^n X_i!} \\ &= \frac{\prod_{i=1}^n Y_i!}{\prod_{i=1}^n X_i!}. \end{aligned} \qquad \text{(0.5 marks)}$$

Since this value is constant with respect to λ, T is a sufficient statistic. If $\bar{X} = \bar{Y}$, then clearly $\sum_{i=1}^n X_i = \sum_{i=1}^n Y_i$, so the ratio of the likelihood functions is the same constant (0.25 marks). This implies T/n is also a sufficient statistic.

Alternatively, note that the parts of $L(\lambda|X_1, \dots, X_n)$ which contain λ depend on X only through T, and use the factorization theorem.

c) (1) Find the maximum likelihood estimator for λ.

Solution. First, notice that $\sum_{i=1}^{n} X_i = n\bar{X}$. Then, the log-likelihood is

$$\begin{aligned}
\ell(\lambda \mid X_1, \ldots, X_n) &= \log L(\lambda \mid X_1, \ldots, X_n) \\
&= \log\left(\frac{\lambda^{\sum_{i=1}^{n} X_i} e^{-n\lambda}}{\prod_{i=1}^{n} X_i!}\right) \\
&= n\bar{X}\log(\lambda) - n\lambda - \log\left(\prod_{i=1}^{n} X_i!\right). \qquad \text{(0.25 marks)}
\end{aligned}$$

Differentiating, the score function is

$$S(\lambda \mid X_1, \ldots, X_n) = \frac{\partial \ell(\lambda \mid X_1, \ldots, X_n)}{\partial \lambda} = \frac{n\bar{X}}{\lambda} - n. \qquad \text{(0.25 marks)}$$

The MLE is the value $\widehat{\lambda}$ such that $S(\widehat{\lambda} \mid X_1, \ldots, X_n) = 0$. Thus,

$$\frac{n\bar{X}}{\widehat{\lambda}} - n = 0 \implies \widehat{\lambda} = \bar{X}. \qquad \text{(0.5 marks)}$$

d) (1) Find a random interval (L_n, U_n) such that $\mathbb{P}(\lambda \in (L_n, U_n)) \approx 1 - \alpha$ for some $\alpha \in (0, 1)$.

Solution. Recall that for an MLE under regularity conditions with $I(\lambda)$ continuous, an approximate α-confidence interval is given by

$$\widehat{\lambda} \pm \frac{z_{\alpha/2}}{\sqrt{I(\widehat{\lambda})}},$$

where $z_{\alpha/2}$ is the $\alpha/2$-quantile of the standard Normal distribution. (0.5 marks)
So, it remains to calculate the observed Fisher information.

$$I(\widehat{\lambda}) = -\frac{\partial^2 \ell(\lambda \mid X_1, \ldots, X_n)}{\partial \lambda^2}\bigg|_{\lambda=\widehat{\lambda}} = -\frac{\partial}{\partial \lambda}\left(\frac{n\bar{X}}{\lambda} - n\right)\bigg|_{\lambda=\widehat{\lambda}} = \frac{n\bar{X}}{\widehat{\lambda}^2} = \frac{n}{\bar{X}}.$$

Thus, the random interval desired is

$$\bar{X} \pm z_{\alpha/2}\sqrt{\frac{\bar{X}}{n}}. \qquad \text{(0.5 marks)}$$

This is great, it's also acceptable for them to use the CLT for IID sums, since the MLE happens to be the sample mean. I believe this gives the same answer, since the variance of a poisson is the mean.

3. *Advanced, 2 marks.* Suppose I flip a coin 10 times and get 7 heads. I claim that the coin is fair. Find an expression for the ~~approximate~~ probability of seeing data at least this extreme if my claim is true. For full marks, show all work and justify all steps using theorems/results from lecture. The density of a binomial random variable is ...

Solution. First, we need to choose a model. Let's suppose that each coin flip is an IID Bernoulli(p) trial, where p is the probability of a head (denoted by an outcome of 1) (0.5 marks). We wish to find the probability of seeing at least 7 heads when $p = 1/2$. By the symmetry of this choice, we must also include the probability of seeing 3 or fewer heads (0.5 marks). Let X denote the sum of the individual flips, so that X is in fact the number of heads observed, and recall that then X has a Binomial(10,1/2) distribution. That is, the probability of interest is

$$\begin{aligned} 1 - \sum_{x=4}^{6} P(X = x) &= 1 - \sum_{x=4}^{6} \binom{10}{x} (1/2)^x (1 - 1/2)^{10-x} && \text{(0.5 marks)} \\ &= 1 - (1/2)^{10} \left[\binom{10}{4} + \binom{10}{5} + \binom{10}{6} \right] \\ &= 0.3438. && \text{(0.5 marks)} \end{aligned}$$

While you are not required to use this terminology in your solution, notice that what you have calculated is a p-value.

THIS PAGE IS FOR ROUGH WORK. NOTHING ON THIS PAGE WILL BE MARKED.

THIS PAGE IS FOR ROUGH WORK. NOTHING ON THIS PAGE WILL BE MARKED.