STAT 415 — Midterm 2 (Spring 2020)

Name: Solutions

Exam rules:

- You have 2hours 20 minutes to complete the exam, scan it and upload onto Ecampus.
- There are **4** Questions.
- This exam is open book. You are free to use your class notes and HW solutions to solve the problems.

  Do not do a brain dump. Answers which are irrelevant will be penalized.

  Do not blindly copy answers from the HW solutions.
- State precisely in the derivations all the results that you use.
- Explain your arguments carefully, show all your work, and write out all your steps. Credit will be given for clarity. No credit will be given for a solution without working.
- If you are caught cheating or helping someone to cheat on this exam, you will both receive a grade of zero on the exam and you will be turned in.
- If a problem requires a numerical answer, you may express answer in terms of elementary functions or cdfs (e.g. $\Gamma(r+1)$ or $r!$).

(1) Suppose that $\underline{X} = (X_1, X_2)'$ are jointly normal where

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\right).$$

Let $Y = (X_1 - 3^{-1}X_2)^2$ and $W = (X_1 + X_2)^2$.

Show that Y and W are independent random variables (stating all the results that you use).

Let $U = X_1 - \frac{1}{3}X_2$ and $V = X_1 + X_2$, then $Y = U^2$ and $W = V^2$. If U and V are independent then any one-to-one transformation $g(u)$ and $h(v)$ will be independent.

<u>Objective</u> Show that U and V are independent this will imply Y and W are independent.

(a) Since U and V are jointly normal. Then U and V are independent if and only if $\operatorname{cov}(U, V) = 0$ (this is <u>not</u> true for non-normal rvs).

$$\operatorname{cov}[U, V] = \operatorname{cov}\left[X_1 - \tfrac{1}{3}X_2,\ X_1 + X_2\right]$$

$$= \operatorname{var}(X_1) - \operatorname{cov}(X_1, X_2) - \tfrac{1}{3}\operatorname{cov}(X_2, X_1) - \tfrac{1}{3}\operatorname{var}(X_2)$$

$$= \operatorname{var}(X_1) - \tfrac{1}{3}\operatorname{var}(X_2) = 1 - \tfrac{1}{3} \times 3 = 0$$

$\Rightarrow$ U and V independent $\Rightarrow$ Y and W are independent $\boxtimes$

(2) Suppose that $\{X_i\}_{i=1}^n$ are iid Poisson distributed random variables with probability mass function $f(k;\lambda) = \lambda^k \exp(-\lambda)/k!$. $\Rightarrow \quad P[X=k ; \lambda] = \frac{\lambda^k e^{-\lambda}}{k!}$

(i) Obtain the maximum likelihood estimator of λ. Denote this estimator as $\widehat{\lambda}$

(ii) Obtain the Mean Squared error: $\mathrm{E}[(\widehat{\lambda}-\lambda)^2]$.

Hint: You can use that

$$\mathrm{var}[X] = \lambda.$$

(iii) Obtain the asymptotic sampling distribution of the MLE.

(iv) Construct the approximate 99% confidence interval for λ based on the MLE.

The confidence interval should be "feasible" (something that a practitioner can immediately evaluate given the MLE).

(v) Obtain the Fisher information of λ.

(vi) For a fixed sample size n, make a sketch of the **log-likelihood** function corresponding to a large Fisher information and a small Fisher information. Indicate on your plots which **log-likelihood** will have an MLE with a small variance and which will have a large variance.

(i) $$L_n(\lambda) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$\mathcal{L}_n(\lambda) = \sum_{i=1}^n \{ x_i \log \lambda - \lambda - \log x_i \}$$

$$\frac{\partial \mathcal{L}_n(\lambda)}{\partial \lambda} = \sum_{i=1}^n \left\{ \frac{x_i}{\lambda} - 1 \right\} = 0$$

$$\Rightarrow \quad \hat{\lambda}_{MLE} = \bar{X}$$

(ii) $$E[\hat{\lambda}_{MLE}] = E[\bar{X}] = E(X_i) = \lambda$$

$\Rightarrow$ The MLE is unbiased and

$$E\{(\hat{\lambda}_{MLE} - \lambda)^2\} = \mathrm{var}\{\hat{\lambda}_{MLE}\} = \mathrm{var}[\bar{X}]$$

$$= \frac{\mathrm{var}(X)}{n} = \frac{\lambda}{n}$$

(iii) Using either Theorem 1.1 or 3.2 we have

$$\sqrt{n}\{\hat{\lambda}_{MLE} - \lambda\} \xrightarrow{D} N(0, \lambda)$$

(iv) The 99% CI for λ will be roughly

$$\left[\hat{\lambda}_{MLE} \pm 2.57 \times \sqrt{\frac{\lambda}{n}}\right]$$

unknown $\Rightarrow$ replace with $\hat{\lambda}_{MLE}$.

$$\left[\hat{\lambda}_{MLE} \pm 2.57 \times \sqrt{\frac{\hat{\lambda}_{MLE}}{n}}\right]$$

(v) $\log f(x;\lambda) = x_i \log \lambda - \lambda - x_i!$

$$\frac{\partial \log f}{\partial \lambda} = \frac{x_i}{\lambda} - 1$$

$$\frac{\partial^2 \log f}{\partial \lambda^2} = -\frac{x_i}{\lambda^2}$$

$$\Rightarrow E\left\{-\frac{\partial^2 \log f}{\partial \lambda^2}\right\} = \frac{E(x_i)}{\lambda^2} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

$\Rightarrow$ Fisher information $= I(\lambda) = \dfrac{1}{\lambda}$

(vi) Observe

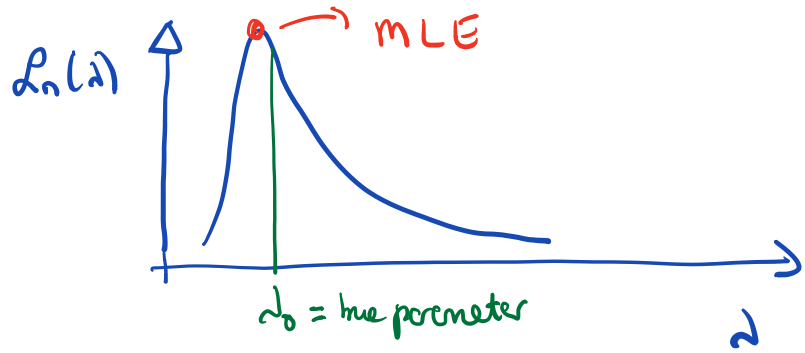
(a) λ = small $\quad I(\lambda)$ large

$\Rightarrow$ curvature of log-likelihood large

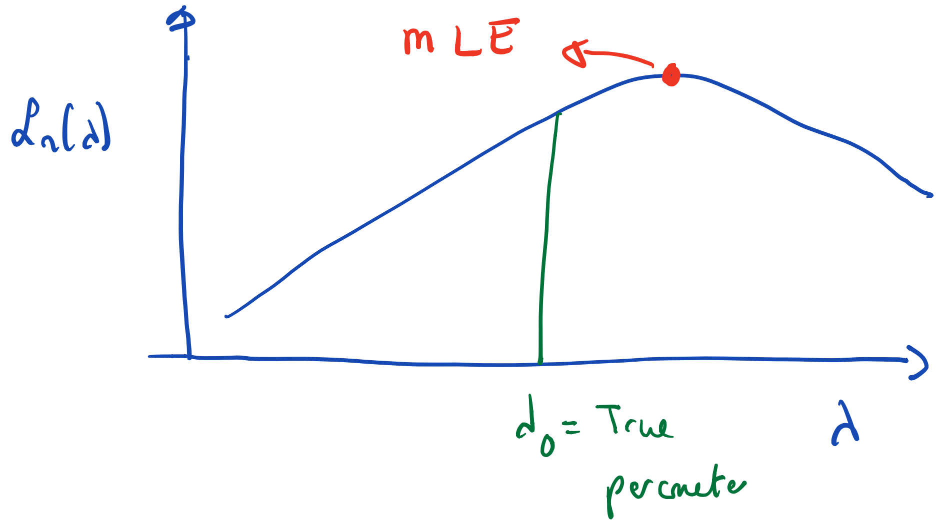
(b) λ = large $I(\lambda)$ small

$\Rightarrow$ curvature of likelihood small.

(a) λ = small



(b) λ = large



(3) Suppose that $\{X_i\}$ are iid random variables with density

$$f(x;\beta) = \beta(1-x)^{\beta-1} \qquad x \in [0,1] \qquad \beta > 0.$$

(i) Evaluate the expectation $\mathrm{E}[X_i]$.
Hint: You may use the two identities

$$\int_0^1 x^a(1-x)^b dx = \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)} \text{ and } \frac{\Gamma(a+1)}{\Gamma(a)} = a,$$

where $\Gamma(a)$ is the Gamma function. If a is an integer, then $\Gamma(a) = (a-1)!$ and $\Gamma(1) = 1$.

(ii) Based on your answer in (i) obtain the method of moments estimator of β. Call this estimator $\widehat{\beta}_{MoM}$.

(iii) Obtain the asymptotic sampling properties of $\widehat{\beta}_{MoM}$.
Hint: You can use that

$$\mathrm{var}[X] = \frac{\beta}{(1+\beta)^2(2+\beta)}$$

(iii) Obtain the maximum likelihood estimator of β. Call this estimator $\widehat{\beta}_{MLE}$.

(iv) Obtain the asymptotic sampling properties of $\widehat{\beta}_{MLE}$.

(v) By comparing the asymptotic variance of $\widehat{\beta}_{MoM}$ and $\widehat{\beta}_{MLE}$ explain for what values of β the two estimators have a variance which is very close.

(i) $$E(X) = \int_0^\infty \beta x(1-x)^{\beta-1} dx$$

$$\boxed{\begin{array}{l} a = 1 \\ b = \beta - 1 \end{array}} \qquad = \beta \int_0^\infty x(1-x)^{\beta-1} dx = \beta \times \frac{\Gamma(1+1)\Gamma(\beta)}{\Gamma(\beta+2)}$$

$$= \beta \times \frac{\Gamma(2)\,\Gamma(\beta)}{\Gamma(\beta+2)} = \beta \times 1 \times \frac{\Gamma(\beta)}{\Gamma(\beta+1)} \times \frac{\Gamma(\beta+1)}{\Gamma(\beta+2)}$$

$$= \frac{\beta}{\beta} \times \frac{1}{\beta+1} = \frac{1}{\beta+1}.$$

(ii) We estimate $E(X)$ with the sample mean $\bar{X}$

Thus the mom estimator of β solves

$$\bar{X} = \frac{1}{1+\hat{\beta}_{mom}} \Rightarrow \hat{\beta}_{mom} = \frac{1-\bar{X}}{\bar{X}}.$$

$$\bar{X}(1+\beta) = 1$$

$$\beta = \frac{1-\bar{X}}{\bar{X}}$$

(iii) we apply Lemma 1.1.

First the regular CLT;

$$\sqrt{n}\left(\bar{X} - \frac{1}{1+\beta}\right) \xrightarrow{P} N\left(0, \frac{\beta}{(1+\beta)^2(2+\beta)}\right).$$

Now by Lemma 1.1 we use that

large n

$$\bar{X} \sim N\left(\frac{1}{1+\beta}, \frac{1}{n(1+\beta)^2(2+\beta)}\right)$$

$$\hat{\beta}_{mom} = g(\bar{X})$$

$$= \left(\frac{1}{\bar{X}} - 1\right)$$

$$g'(\mu) = -\frac{1}{\mu^2} \Rightarrow g'(\mu) = -(1+\beta)^2$$

$$\Rightarrow [g'(\mu)]^2 = (1+\beta)^4$$

Thus by Lemma 1.1 we have

$$\sqrt{n}\{\hat{\beta}_{mom} - \beta\} \longrightarrow N\left(0, \frac{\beta(1+\beta)^2}{(2+\beta)}\right)$$

For large n

$$\hat{\beta}_{mom} \sim N\left(\beta, \frac{\beta(1+\beta)^2}{n(2+\beta)}\right)$$

(iii) <u>The MLE;</u>

$$L_n(\beta) = \prod_{i=1}^{n} \beta (1-x_i)^{\beta-1}$$

$$\mathcal{L}_n(\beta) = \sum_{i=1}^{n} \left\{ \log \beta + (\beta-1) \log (1-x_i) \right\}$$

$$\frac{\partial \mathcal{L}_n}{\partial \beta} = \sum_{i=1}^{n} \left\{ \frac{1}{\beta} + \log (1-x_i) \right\}$$

Solving $\frac{\partial \mathcal{L}_n}{\partial \beta} = 0$ gives the MLE

$$\hat{\beta} = \left(- \frac{1}{n} \sum_{i=1}^{n} \log (1-x_i) \right)^{-1}$$

(iv) To obtain the Fisher information:

$$\frac{\partial^2 \log f}{\partial \beta^2} = - \frac{1}{\beta^2}$$

$$\Rightarrow I(\beta) = - E \left\{ \frac{\partial^2 \log f}{\partial \beta^2} \right\} = \frac{1}{\beta^2}$$

Thus by Theorem 3.2 we have

$$\sqrt{n} \left[\hat{\beta}_{MLE} - \beta \right] \rightarrow N(0, \beta^2)$$

For large n

$\hat{\beta} \sim N\left(\beta, \frac{\beta^2}{n}\right)$

(v) The asymptotic variances are

$\boxed{\text{MoM}}$ $\quad \dfrac{\beta(1+\beta)^2}{n(2+\beta)}$ $\qquad$ $\boxed{\text{MLE}}$ $\quad \dfrac{\beta^2}{n}$

We know that the MLE has a smaller variance. Our aim is to deduce when the two estimators have a variance which is very close.

Since

$$\underbrace{\frac{\beta(1+\beta)^2}{(2+\beta)}}_{\text{MoM}} > \underbrace{\beta^2}_{\text{MLE}}$$

$$\frac{(1+\beta)^2}{\beta(2+\beta)} > 1$$

$$\Rightarrow \quad 1 + \frac{1}{\beta^2+2\beta} > 1.$$

But we observe if β is large, then the variance of the MoM and MLE will be close, since large $\beta \Rightarrow \dfrac{1}{\beta^2+2\beta}$ = very small.

(4) Suppose that $\{X_i\}_{i=1}^n$ are iid random variables with density

$$f(x;\alpha) = \begin{cases} \frac{1}{10-\alpha} & x \in [10-\alpha, 10] \\ 0 & x \notin [10-\alpha, 10] \end{cases}$$

Obtain the maximum likelihood estimator of α.

I made a mistake! This should read $\frac{1}{\alpha}$, else it does not integrate to 1

The likelihood is

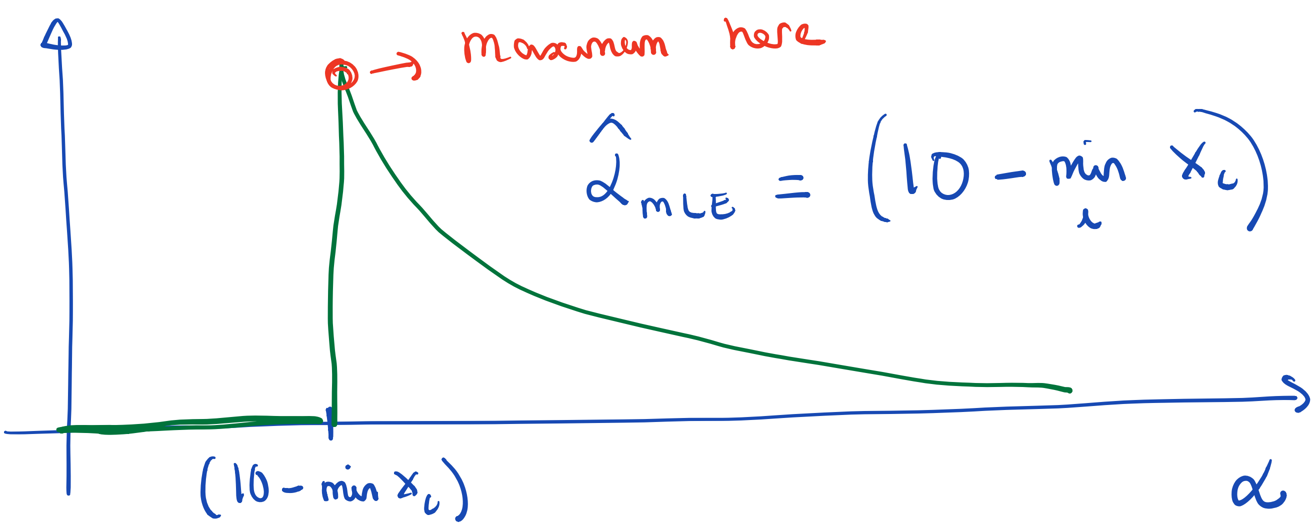
$$L_n(\alpha) = \prod_{i=1}^{n} \frac{1}{\alpha} \underbrace{I_{[10-\alpha, 10]}(X_i)}_{10-\alpha < X_i \text{ for all } i.}$$

Else it is zero.

$$= \frac{1}{\alpha^n} \underbrace{I_{[10-\alpha, 10]}(\min_i X_i)}_{10-\alpha < \min X_i}$$

$$\Rightarrow \alpha > 10 - \min X_i$$

else it is zero.



$$\hat{\alpha}_{MLE} = (10 - \min_i X_i)$$