

MATH 368/621 Fall 2020 Homework #6

Professor Adam Kapelner

Due by email 11:59PM, Monday, Nov 30, 2020

(this document last updated Sunday 22nd November, 2020 at 11:51am)

Instructions and Philosophy

The path to success in this class is to do many problems. Unlike other courses, exclusively doing reading(s) will not help. Coming to lecture is akin to watching workout videos; thinking about and solving problems on your own is the actual “working out.” Feel free to “work out” with others; **I want you to work on this in groups.**

Reading is still *required*. For this homework set, read on your own about the sample variance rv, quadratic forms, Cochran’s theorem and the multivariate normal.

The problems below are color coded: **green** problems are considered *easy* and marked “[easy]”; **yellow** problems are considered *intermediate* and marked “[harder]”, **red** problems are considered *difficult* and marked “[difficult]” and **purple** problems are extra credit. The *easy* problems are intended to be “giveaways” if you went to class. Do as much as you can of the others; I expect you to at least attempt the *difficult* problems.

This homework is worth 100 points but the point distribution will not be determined until after the due date. See syllabus for the policy on late homework.

Up to 7 points are given as a bonus if the homework is typed using L^AT_EX. Links to installing L^AT_EX and program for compiling L^AT_EX is found on the syllabus. You are encouraged to use **overleaf.com**. If you are handing in homework this way, read the comments in the code; there are two lines to comment out and you should replace my name with yours and write your section. The easiest way to use overleaf is to copy the raw text from hwxx.tex and preamble.tex into two new overleaf tex files with the same name. If you are asked to make drawings, you can take a picture of your handwritten drawing and insert them as figures or leave space using the “\vspace” command and draw them in after printing or attach them stapled.

The document is available with spaces for you to write your answers. If not using L^AT_EX, print this document and write in your answers. I do not accept homeworks which are *not* on this printout. Keep this first page printed for your records.

NAME: _____ SECTION: _____ CLASS: 368 | 621

Problem 1

The χ^2 r.v. within Cochran's Theorem.

- (a) [easy] Given $X_1, \dots, X_n \stackrel{iid}{\sim} f(\mu, \sigma^2)$, a density with finite variance, state the classic estimator S^2 (a r.v.) and the estimate (a scalar value) for σ^2 , the variance of the X 's.

- (b) [difficult] [MA] Prove $\mathbb{E}[\cdot] = \sigma^2$. The answer is online but try to do it yourself. (This property is called “unbiasedness” in a statistical inference context.

- (c) [easy] Given $X_1, \dots, X_n \stackrel{iid}{\sim} f(\mu, \sigma^2)$, a density with finite variance, state the classic estimator S (a r.v.) and the estimate (a scalar value) for σ , the standard error of the X 's.

(d) [E.C.] Prove this estimator is *biased* i.e $\mathbb{E}[\cdot] \neq \sigma$.

(e) [easy] State Cochran's Theorem.

(f) [easy] Given $X_1, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$. Show that $\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_n^2$.

- (g) [easy] Let $Z_1 := \frac{X_1 - \mu}{\sigma}, \dots, Z_n := \frac{X_n - \mu}{\sigma}$. We know that $Z_1, \dots, Z_n \stackrel{iid}{\sim} \mathcal{N}(0, 1)$ and let the column vector r.v. $\mathbf{Z} := [Z_1 \ \dots \ Z_n]^\top$. Express $\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2$ in vector notation using \mathbf{Z} .
- (h) [easy] Express $\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2$ as a quadratic form. What is the matrix that determines this quadratic form?
- (i) [easy] What is the rank of the determining matrix?
- (j) [easy] When computing $\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2$, how many independent pieces of information AKA “degrees of freedom” go into the calculation?
- (k) [easy] Show that $\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 = \frac{(n-1)S^2}{\sigma^2} + \frac{n(\bar{X} - \mu)^2}{\sigma^2}$.
- (l) [easy] Show that $\frac{n(\bar{X} - \mu)^2}{\sigma^2} \sim \chi_1^2$.

(m) [easy] Express $\frac{n(\bar{X}-\mu)^2}{\sigma^2}$ as a quadratic form. What is the matrix that determines this quadratic form? Call it B_2 .

(n) [easy] What is the rank of the determining matrix?

(o) [easy] Express $\frac{(n-1)S^2}{\sigma^2}$ in vector notation.

(p) [harder] Express $\frac{(n-1)S^2}{\sigma^2}$ as a quadratic form. What is the matrix that determines this quadratic form? Call it B_1 .

(q) [harder] What is the rank of the determining matrix?

(r) [easy] When computing $\frac{(n-1)S^2}{\sigma^2}$, how many independent pieces of information go into the calculation?

(s) [easy] What is $B_1 + B_2$?

(t) [easy] What is $\text{rank}(B_1) + \text{rank}(B_2)$?

(u) [easy] Are the conditions of Cochran's Theorem satisfied so that we can conclude that $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$ and that $\frac{(n-1)S^2}{\sigma^2}$ is independent of $\frac{n(\bar{X}-\mu)^2}{\sigma^2}$? Yes or no.

(v) [E.C.] Prove Cochran's Theorem. Do on a separate sheet.

(w) [difficult] [MA] What is $B_1 B_2$? Why do you think this should be?

(x) [harder] Using your previous answers, show that $\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} \sim T_{n-1}$.

(y) [easy] Make up a definition of "degrees of freedom" in English.

(z) [harder] What is the distribution of S^2 ?

Problem 2

More vector r.v. operations.

(a) [harder] Let \mathbf{X} be a vector r.v. with dimension n , $\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu}$ and $\text{Var}[\mathbf{X}] = \Sigma$. Let $A \in \mathbb{R}^{m \times n}$ be a matrix of constants. Show that $\mathbb{E}[A\mathbf{X}] = A\boldsymbol{\mu}$.

(b) [harder] Show that $\text{Var}[A\mathbf{X}] = A\Sigma A^\top$.

Problem 3

Some questions about ch.f.'s and the MVN really quickly

- (a) [easy] Let $Z_1, \dots, Z_n \stackrel{iid}{\sim} \mathcal{N}(0, 1)$ and let the column vector r.v. $\mathbf{Z} := [Z_1 \ \dots \ Z_n]^\top$. What is the PDF of \mathbf{Z} ? How is it distributed?

- (b) [harder] Find $\phi_{\mathbf{Z}}(\mathbf{t})$. Remember \mathbf{t} is a column vector of dimension n .

- (c) [easy] Let $A \in \mathbb{R}^{n \times n}$ be an invertible matrix of constants and $\boldsymbol{\mu} \in \mathbb{R}^n$ be a vector of constants. What is the PDF of $\mathbf{X} = A\mathbf{Z} + \boldsymbol{\mu}$? It should be a function of $n, \mathbf{x}, \boldsymbol{\mu}, A$ and fundamental constants only.

- (d) [difficult] [MA] Prove this PDF by using change of variables.

- (e) [easy] Let $A \in \mathbb{R}^{m \times n}$ be a non-square matrix of constants and $\boldsymbol{\mu} \in \mathbb{R}^n$ be a vector of constants. What is the PDF of $\mathbf{X} = A\mathbf{Z} + \boldsymbol{\mu}$? It should be a function of $n, \mathbf{x}, \boldsymbol{\mu}, A$ and fundamental constants only.
- (f) [difficult] [MA] Prove this PDF. You cannot use change of variables since $\mathbf{X} = g(\mathbf{Z})$ is not invertible. You will have to use another method.
- (g) [harder] Let $A \in \mathbb{R}^{m \times n}$ be a non-square matrix of constants and $\boldsymbol{\mu} \in \mathbb{R}^n$ be a vector of constants. What is the ch.f. of $\mathbf{X} = A\mathbf{Z} + \boldsymbol{\mu}$? It should be a function of $n, \mathbf{t}, \boldsymbol{\mu}, A$ and fundamental constants only. Use property 2 of multivariate ch.f.'s.
- (h) [difficult] Show that $(\mathbf{X} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_n^2$. This amounts to repeating a proof from class.

(i) [harder] Let $\mathbf{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \Sigma)$. Let $B \in \mathbb{R}^{m \times n}$ be a matrix of constants and $\mathbf{c} \in \mathbb{R}^m$ be a vector of constants. Find the distribution of $\mathbf{Y} = B\mathbf{X} + \mathbf{c}$.

(j) [harder] [MA] Let $\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{iid}{\sim} \mathcal{N}_n(\boldsymbol{\mu}, \Sigma)$. Using properties 2 and 3 of ch.f.'s, find the distribution of the average $\bar{\mathbf{X}} = \frac{1}{n}(\mathbf{X}_1 + \dots + \mathbf{X}_n)$.

(k) [difficult] [MA] Let $\mathbf{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \Sigma)$ where $n > 5$. Find the distribution of the first five components, $\mathbf{U} = [X_1 \ X_2 \ X_3 \ X_4 \ X_5]^\top$.