MATH 368/621 Fall 2019 Homework #6

Professor Adam Kapelner

Due under the door of KY604 11:59PM Monday, December 2, 2019

(this document last updated Thursday 21st November, 2019 at 6:27pm)

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 central limit theorem, the normal, χ^2_k, F, T, cauchy, quadratic forms, matrix rank, the estimator S^2 for sample variance, Cochran's Theorem, degrees of freedom, multivariate normal distribution and multivariate characteristics functions.

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 LaTeX. Links to instaling LaTeX and program for compiling LaTeX 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 LaTeX, 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

Introducing the king: the normal distribution $\mathcal{N}$ and his princes/sses: the lognormal distribution Log$\mathcal{N}$, chi-squared distribution χ^2_k, Student's T distribution T_k and Fisher-Snecodor's distribution F_{k_1,k_2}.

(a) [easy] Let $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ independent of $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$. Prove $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ using ch.f.'s.

(b) [E.C.] Let $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ independent of $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$. Prove $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ using the definition of convolution on a separate page. This is in the book but try not to look at it.

(c) [easy] Let $X \sim \text{Log}\mathcal{N}(\mu, \sigma^2)$ and $Y = \ln(X)$. How is Y distributed? Use a heuristic argument. No need to actually change variables.

(d) [harder] Let $X_1 \sim \text{Log}\mathcal{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \text{Log}\mathcal{N}(\mu_2, \sigma_2^2), \ldots, X_n \sim \text{Log}\mathcal{N}(\mu_n, \sigma_n^2)$ all independent of each other and $Y = \prod_{i=1}^n X_i$. How is Y distributed? Use a heuristic argument. No need to actually change variables.

(e) [easy] Let $X \sim \chi^2_k$, find $\mathbb{E}[X]$ using the fact that $X = Z_1^2 + Z_2^2 + \ldots + Z_k^2$ where $Z_1, Z_2, \ldots, Z_k \overset{iid}{\sim} \mathcal{N}(0, 1)$.

(f) [easy] Let $X \sim \chi^2_k = \text{Gamma}\left(\frac{k}{2}, \frac{1}{2}\right)$. Find the PDF of X by making the correct substitutions in the gamma PDF and simplifying.

(g) [easy] Using $Z_1, Z_2, \ldots \overset{iid}{\sim} \mathcal{N}(0, 1)$, the function g s.t. $g(Z_1, Z_2, \ldots) \sim \chi^2_k$ where $k \in \mathbb{N}$ is a constant is given below:

$$g(Z_1, Z_2, \ldots) = Z_1^2 + Z_2^2 + \ldots + Z_k^2 \sim \chi^2_k$$

Following this example, find a function g s.t. $g(Z_1, Z_2, \ldots) \sim F_{k_1,k_2}$ where $k_1, k_2 \in \mathbb{N}$ are constants.

(h) [easy] Let $X \sim F_{k_1,k_2}$, find the kernel of $f_X(x)$.

(i) [easy] Using $Z_1, Z_2, \ldots \overset{iid}{\sim} \mathcal{N}(0, 1)$, find a function g s.t. $g(Z_1, Z_2, \ldots) \sim T_k$ where $k \in \mathbb{N}$ is a constant.

(j) [easy] Let $X \sim T_k$, find the kernel of $f_X(x)$.

(k) [E.C.] Derive the PDF of the T_k distribution using the ratio formula where you first find the distribution of the denominator explicitly. Do on a separate piece of paper.

(l) [E.C.] Show that the PDF of $X \sim T_k$, converges to the PDF of $Z \sim \mathcal{N}(0, 1)$ when $k \to \infty$. Hint: use Stirling's approximation. Do on a separate piece of paper.

(m) [easy] Let $X \sim \text{Cauchy}(0, 1)$, find the kernel of $f_X(x)$.

(n) [easy] Using $Z_1, Z_2, \ldots \overset{iid}{\sim} \mathcal{N}(0, 1)$, find a function g s.t. $g(Z_1, Z_2, \ldots) \sim \text{Cauchy}(0, 1)$.

(o) [easy] Let $X \sim \text{Cauchy}(0, 1)$, prove that $\mathbb{E}[X]$ does not exist without using its ch.f.

Problem 2

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

(a) [easy] Given $X_1, \ldots, X_n \overset{iid}{\sim} f(\mu, \sigma^2)$, 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 this estimator is unbiased i.e $\mathbb{E}[\cdot] = \sigma^2$. The answer is online but try to do it yourself.

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

(g) [easy] Let $Z_1 := \frac{X_1 - \mu}{\sigma}, \ldots, Z_n := \frac{X_n - \mu}{\sigma}$. We know that $Z_1, \ldots, Z_n \overset{iid}{\sim} \mathcal{N}(0,\, 1)$ and let the column vector r.v. $\boldsymbol{Z} := [Z_1 \ \ldots \ Z_n]^\top$. Express $\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2$ in vector notation using $\boldsymbol{Z}$.

(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 $\frac{n(\bar{X} - \mu)^2}{\sigma^2} \sim \chi^2_1$.

(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 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ 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".

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

(aa) [difficult] [MA] What is $\mathbb{E}[S]$?

(bb) [difficult] [MA] Create a new estimator S_0 that is unbiased for σ i.e. ($\mathbb{E}[S] = \sigma$). Hint: use S but multiply by intelligent constants.

Problem 3

More vector r.v. operations.

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

(b) [harder] Show that $\mathbb{V}\mathrm{ar}\left[A\boldsymbol{X}\right] = A\Sigma A^\top$.

Problem 4

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

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

(b) [harder] Find $\phi_{\boldsymbol{Z}}(\boldsymbol{t})$. Remember $\boldsymbol{t}$ 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 $\boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu}$? It should be a function of $n, \boldsymbol{x}, \boldsymbol{\mu}, A$ 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 $\boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu}$? It should be a function of $n, \boldsymbol{x}, \boldsymbol{\mu}, A$ and fundamental constants only.

(f) [difficult] [MA] Prove this PDF. You cannot use change of variables since $\boldsymbol{X} = g(\boldsymbol{Z})$ 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 $\boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu}$? It should be a function of $n, \boldsymbol{t}, \boldsymbol{\mu}, A$ and fundamental constants only. Use property 2 of multivariate ch.f.'s.

(h) [difficult] Show that $(\boldsymbol{X} - \boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2_n$. This amounts to repeating a proof from class.

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

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

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

Problem 5

That job interview question...

(a) [E.C.] Consider a circle. Place 3 points inside the circle at random. Connect the three points by lines to form a triangle. What is the probability the triangle contains the center of the circle? Have fun...