MATH 340/640 Fall 2025 Homework #5

Professor Adam Kapelner

Due by email on the date found on the homepage

(this document last updated Tuesday 18$^{\text{th}}$ November, 2025 at 6:05pm)

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* — read about the concepts we discussed in class online. For this homework set, also review Math 241 notes about the normal distribution.

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. "[MA]" are for those registered for 621 and extra credit otherwise.

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 5 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: ____

Problem 1

These exercises will give you practice with the gamma function.

(a) [in the notes] Write the PMF's and parameter spaces of both the extended negative binomial rv and the negative binomial rv model. Explain how the latter "upgrades" the former.

Problem 2

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

(a) [harder] Let $X \sim \text{Log}\mathcal{N}(\mu, \sigma^2)$ and $Y = \ln(X)$. Prove that $Y \sim \mathcal{N}(\mu, \sigma^2)$.

(b) [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.

(c) [difficult] [MA] 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. This is a lot of boring algebra but it will hone your skills. You can find it in the book or on the Internet (but try not to look at the answer).

(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. Hint: how are the e^{X_i}'s distributed?

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

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

As this series of questions caused confusion in previous years, I will give you the first answer here. This answer then serves as a model answer for the next three problems as the answer will always be in this type of style.

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

(g) [harder] 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.

(h) [harder] Using $Z_1, Z_2, \ldots \overset{iid}{\sim} \mathcal{N}(0, 1)$, find a function g s.t. $g(Z_1, Z_2, \ldots) \sim F_{k_1,k_2}$ where $k_1, k_2 \in \mathbb{N}$ are constants.

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

(j) [difficult] If $W^2 \sim F_{1,k}$ $k \in \mathbb{N}$, W is symmetric about zero, then how is W distributed? Hint: $F_{W^2}(w) := P(W^2 \leq w) = P(W \in [-\sqrt{w}, \sqrt{w}]) = F_W(\sqrt{w}) - F_W(-\sqrt{w})$ then take the derivative wrt w on both sides.

(k) [difficult] Assume W to be the answer from the previous question. Let $X = g(W) = |W|$. Find the PDF of X. Remember that g is not a 1:1 function. Hint: $F_X(x) := P(X \leq x) = P(|W| \leq x) = P(W \in [-x, x])\mathbb{1}_{x>0} = (F_W(x) - F_W(-x))\mathbb{1}_{x>0}$. Then take the derivative wrt x on both sides.

(l) [difficult] Let X_n be a sequence of continuous rv's. Prove that if $\lim_{n\to\infty} f_{X_n}(x) = f_X(x)$ then $X_n \overset{d}{\to} X$ i.e., PDF convergence implies CDF convergence. Hint: in our case here, assume you can interchange limits and integrals as $f_{X_n}(x)$ doesn't do weird stuff like shoot off to infinity.

(m) [E.C.] Show that the PDF of $X_n \sim T_n$, converges to the PDF of $Z \sim \mathcal{N}(0, 1)$ when $n \to \infty$. Using the previous fact, this demonstrates that the T distribution converges to the normal distribution as the degrees of freedom get large, an important fact in statistics. This convergence is so rapid that usually at even at $n \geq 30$, people just use the normal distribution. Hint: use Stirling's approximation.

(n) [difficult] Let $X \sim F_{a,b}$. Show that $\frac{a}{b}X \sim \text{BetaPrime}(a/2, b/2)$ where $a, b > 0$.

(o) [difficult] Let $X \sim \text{Cauchy}(0, 1)$. Prove that $\mu := \mathbb{E}[X]$ does not exist (thus the WLLN does not apply) using chf's. (In class, we proved this a different way using the definition of the doubly improper integral).

(p) [in the notes] Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Cauchy}(0, 1)$. Using chf's, show that the distribution of $\bar{X}_n$ does not converge in distribution to a constant.

Problem 3

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

(a) [easy] State Cochran's Theorem (the assumptions and the equivalent results).

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

(c) [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 $\boldsymbol{Z}$.

(d) [harder] 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? (This is the matrix sandwiched between the two vectors).

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

(f) [harder] 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?

(g) [easy] Show that $\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 = \frac{(n-1)S^2}{\sigma^2} + \frac{n(\bar{X}-\mu)^2}{\sigma^2}$.

(h) [easy] Show that $\frac{n(\bar{X}-\mu)^2}{\sigma^2} \sim \chi^2_1$.

(i) [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.

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

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

(l) [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.

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

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

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

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

(q) [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.

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

(s) [easy] Write a definition of "degrees of freedom" in English.

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

Problem 4

We will now practice multivariate change of variables where $\boldsymbol{Y} = \boldsymbol{g}(\boldsymbol{X})$ where $\boldsymbol{X}$ de a vector of k continuous r.v.'s and $\boldsymbol{g} : \mathbb{R}^k \to \mathbb{R}^k$ and is 1:1.

(a) [easy] State the formula for the PDF of $\boldsymbol{Y}$.

(b) [harder] Demonstrate that the formula for the PDF of $\boldsymbol{Y}$ reduces to the univariate change of variables formula if the dimensions of $\boldsymbol{Y}$ and $\boldsymbol{X}$ are 1.

(c) [easy] State the integral formula for the PDF of $R = \frac{X_1}{X_2}$ if X_1 and X_2 are dependent.

(d) [easy] State the integral formula for the PDF of $R = \frac{X_1}{X_2}$ if X_1 and X_2 are iid.

(e) [harder] State the integral formula for the PDF of $R = \frac{X_1}{X_2}$ if X_1 and X_2 are iid and have positive supports.

(f) [easy] Show that $R = \frac{X_1}{X_2} \sim \beta'(\alpha, \beta)$, the beta prime distribution, if $X_1 \sim \text{Gamma}\,(\alpha, 1)$ independent of $X_2 \sim \text{Gamma}\,(\beta, 1)$. Be careful to include the indicator function for r in the final result.

(g) [in the notes] Put it all together: let $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}\,(\mu, \sigma^2)$. Derive the PDF of $R = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$ from scratch using Cochran's theorem and the integral formula for the ratio.

(h) [difficult] Find an integral formula for the PDF of $E = X_1^{X_2}$ if X_1 and X_2 are dependent.

(i) [difficult] Find an integral formula for the PDF of $R = \frac{X_1}{X_1+X_2}$ if X_1 and X_2 are dependent.

(j) [easy] State the integral formula for the PDF of $R = \frac{X_1}{X_1+X_2}$ if X_1 and X_2 are independent.

(k) [easy] State the integral formula for the PDF of $R = \frac{X_1}{X_1+X_2}$ if X_1 and X_2 are iid.

(l) [harder] Find an integral formula for the PDF of $R = \frac{X_1}{X_1+X_2}$ if X_1 and X_2 are independent and have positive supports.

(m) [difficult] [MA] Find the simplest integral formula you can for the PDF of $Q = \frac{X_1}{X_2}e^{X_3}$ where X_1, X_2, X_3 are dependent continuous r.v.'s.