MATH 340/640 Fall 2024 Homework #6

Professor Adam Kapelner

Due by email 11:59PM November 25, 2024

(this document last updated Tuesday 19^{th} November, 2024 at 8:57pm)

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, review the previous random variables and read about mixture/compound distributions and order statistics.

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 640 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

We will practice finding kernels and relating them to known distributions. The gamma function and the beta function will come up as well.

(a) [in the notes] Find the kernel of $X \sim \text{Binomial}\,(n,\, p)$.

(b) [in the notes] Find the kernel of $X \sim \text{Mult}_k\,(n, \boldsymbol{p})$.

(c) [easy] Find the kernel of $X \sim \text{Gamma}\,(\alpha,\, \beta)$.

(d) [easy] Find the kernel of $X \sim \text{BetaPrime}\,(\alpha, \beta)$.

(e) [easy] Find the kernel of $\text{Cauchy}\,(\mu, \sigma)$.

(f) [easy] Find the kernel of the Lomax rv's PDF. See its Wikipedia article here.

(g) [easy] Find the kernel of the PDF of $X \sim \text{Log}\mathcal{N}(\mu, \sigma^2)$.

(h) [harder] If $k(x) = e^{-\lambda x} x^{k-1} \mathbb{1}_{x>0}$ how would you know if the r.v. X was an Erlang (k, λ) or a Gamma (k, λ)?

(i) [harder] If $k(x) = e^{ax - bx^2}$, solve for the normalization constant c.

(j) [harder] If $k(x) = xe^{-x^2} \mathbb{1}_{x>0}$, how is X distributed?

(k) [in the notes] Fill in the integral definition. $B(\alpha, \beta) :=$

(l) [in the notes] Prove $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ using the method from class.

(m) [easy] Find the kernel of the PDF of $X \sim T_\nu$.

Problem 2

We will now practice using order statistics concepts.

(a) [easy] If $X_1, \ldots, X_n \stackrel{iid}{\sim}$ some continous rv where PDF is denoted $f(x)$ and its CDF is denoted $F(x)$, express the CDF of the maximum X_i and express the CDF of the minimum X_i.

(b) [easy] If $X_1, \ldots, X_n \stackrel{iid}{\sim}$ some continous rv where PDF is denoted $f(x)$ and its CDF is denoted $F(x)$, express the PDF of the maximum X_i and express the PDF of the minimum X_i.

(c) [easy] If $X_1, \ldots, X_n \stackrel{iid}{\sim}$ some continous rv where PDF is denoted $f(x)$ and its CDF is denoted $F(x)$, express the *CDF* of $X_{(k)}$ i.e. the kth smallest X_i.

(d) [harder] If $X_1, \ldots, X_n \overset{iid}{\sim}$ some continous rv where its PDF is denoted $f(x)$ and its CDF is denoted $F(x)$, express the kernel of the PDF of $X_{(k)}$ i.e. the kth smallest X_i.

(e) [harder] If discrete $X_1, \ldots, X_n \overset{iid}{\sim}$ some discrete rv, why would the formula in (b) not be correct?

(f) [harder] If $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{U}(0,\, 1)$, show that $X_{(k)} \sim \mathrm{Beta}(k,\, n-k+1)$.

(g) [harder] Express $\binom{n}{k}$ in terms of the beta function.

(h) [easy] Prove that $\text{Beta}(1, 1) = \text{U}(0, 1)$.

(i) [difficult] If $X \sim \text{Binomial}(n, p)$, show that $F(x) = I_{1-p}(n-k, k+1)$. You will need to assume the property $I_x(a, b) = 1 - I_{1-x}(b, a)$.

This answer is done for you. See last page of the homework so you can copy it.

Problem 3

We will practice using computing quantiles, finding the quantile function and sampling.

(a) [in the notes] Let X be a continous rv. Prove $Y = F_X(X) \sim \mathrm{U}\left(0,\, 1\right)$.

(b) [harder] If $X \sim \mathrm{Logistic}(0, 1)$, find $\mathrm{MED}\left[X\right]$.

(c) [easy] Write an algorithm for sampling from $X \sim \mathrm{U}\left(a,\, b\right)$.

(d) [harder] Write an algorithm for sampling from $X \sim \text{Log}\mathcal{N}(\mu,\, \sigma^2)$.

(e) [difficult] Write an algorithm for sampling from $X \sim \text{NegBin}(k, p)$.

Problem 4

We will now practice the conditional-on-total distributions.

(a) [easy] Prove the PMF of $X \sim \text{Poisson}(\lambda)$ using the limit as $n \to \infty$ and let $p = \frac{\lambda}{n}$.

(b) [difficult] Let $X_1, X_2 \overset{iid}{\sim} \text{Geometric}\,(p)$ and $T = X_1 + X_2$. Find the PMF of $X_1 \mid T = t$. It will be a brand name random variable.

(c) [harder] [MA] Let $X_1, X_2 \overset{iid}{\sim} \text{Binomial}\,(n,\, p)$ and $T = X_1 + X_2$. Show that $X_1 \mid T = t$ is hypergeometric. You can find information about this r.v. online. Note we did not / will not study the hypergeometric further and it will not be covered on any exams.

Problem 5

We will now practice with mixture and compound distributions.

(a) [harder] If $X \sim \text{Bernoulli}\,(0.17)$ and $Y \mid X = x \sim \mathcal{N}\,(x,\, 1)$, find the PDF of Y.

(b) [harder] Find the $\mathbb{P}(X = 1 \mid Y = 2)$.

(c) [easy] If $X \sim \text{Beta}(\alpha, \beta)$ and $Y \mid X = x \sim \text{Binomial}(n, x)$, find the PDF of Y.

(d) [difficult] If $X \sim \text{Gamma}(\alpha, \beta)$ and $Y \mid X = x \sim \text{Exp}(x)$, show that $Y \sim \text{Lomax}(\alpha, \beta)$. See the Lomax's Wikipedia article by clicking here. Kernels always make it easier.

Problem 6

We will acquaint ourselves with finding modes of distributions.

(a) [harder] Let $X \sim \text{Gamma}(\alpha, \beta)$. Find Mode$[X]$. Any restrictions on α, β besides the parameter space?

(b) [harder] Let $X \sim \text{InvGamma}(\alpha, \beta) := \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\frac{\beta}{x}} \mathbb{1}_{x \in (0,\infty)}$. Find Mode$[X]$. Any restrictions on α, β besides the parameter space (which is idential to the Gamma distribution's parameter space)?

(c) [harder] Let $X \sim T_k$. Let $Y = \mu + \sigma X$ where $\mu \in \mathbb{R}$ and $\sigma > 0$. We proved on the 2023 midterm II that $Y \sim T_k(\mu, \sigma^2)$ with density:

$$f_Y(y) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k \sigma^2}\,\Gamma\left(\frac{k}{2}\right)} \left(1 + \frac{(y-\mu)^2}{k\sigma^2}\right)^{-\frac{k+1}{2}}$$

which is called the "location-scale T-distribution". Find Mode$[Y]$.

Solution for the Binomial CDF question

Assume $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{U}\,(0,\, 1)$. We know then that $F_X(x) = x$. So by the CDF formula for an order statistic k and the fact that the CDF is a probability we have:

$$\begin{aligned}
F_{X_{(k)}}(x) &= \sum_{j=k}^{n} \binom{n}{j} F_X(x)^j \left(1 - F_X(x)\right)^{n-j} \\
&= \sum_{j=k}^{n} \binom{n}{j} x^j \left(1 - x\right)^{n-j} \\
&= 1 - \sum_{j=0}^{k-1} \binom{n}{j} x^j \left(1 - x\right)^{n-j}
\end{aligned}$$

From class we proved that the order statistics for the standard uniform are distributed beta and we also know its CDF:

$$X_{(k)} \sim \mathrm{Beta}\,(k,\, n-k+1) = \mathrm{Beta}\,(k,\, n-(k-1)) \quad \Rightarrow \quad F_{X_{(k)}}(x) = I_x(k, n-(k-1))$$

By the fact on wikipedia about the I function we have:

$$F_{X_{(k)}}(x) = I_x(k, n-(k-1)) = 1 - I_{1-x}(n-(k-1), k)$$

Thus we have the strange equality:

$$\sum_{j=0}^{k-1} \binom{n}{j} x^j \left(1 - x\right)^{n-j} = I_{1-x}(n-(k-1), k)$$

Letting $y := k - 1$,

$$\sum_{j=0}^{y} \binom{n}{j} x^j \left(1 - x\right)^{n-j} = I_{1-x}(n-y, y+1)$$

Note that the lhs is the CDF for $Y \sim \mathrm{Binomial}\,(n,\, x)$. $\blacksquare$