MATH 368/621 Fall 2019 Homework #4

Professor Adam Kapelner

Due under the door of KY604 11:59PM Friday, November 1, 2019

(this document last updated Friday 8^{th} November, 2019 at 12:10 Noon)

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, review from math 241 about conditional probability, expectation and variance then read on your own about the multinomial distribution, conditional vector expectation, covariances, variance-covariance matrices.

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

These exercises will give you practice with the gamma function.

(a) [easy] Write the definition of $\Gamma(x)$.

(b) [easy] Prove $\Gamma(x+1) = x\Gamma(x)$.

(c) [easy] Write the definition of $\Gamma(x, a)$ without using the gamma function.

(d) [harder] Write the definition of $Q(x, a)$ without using the gamma function.

(e) [harder] If $0 < a < b < \infty$, find an expression for $\Gamma(x, b) - \gamma(x, a)$.

(f) [easy] For $a, c \in (0, \infty)$, prove the following:

$$\int_a^\infty t^{x-1} e^{-ct} dt = \frac{\Gamma(x, ac)}{c^x}$$

(g) [easy] Let $X \sim \text{Gamma}(\alpha, \beta)$. Show that this r.v. is equivalent to $X \sim \text{Erlang}(k, \lambda)$ and find k and λ in terms of α and β. Are there any restrictions on the values of α and β for this relationship to hold?

Problem 2

These exercises will give you practice with the Poisson process and the analogous Binomial-Negative Binomial relationship.

(a) [easy] Write the assumptions and the main result of the Poisson process (an equivalence of two probability statements and then an equivalence using the CDF's of the Erlang and the Poisson models).

Let

$$\begin{aligned} T_k &\sim \text{Erlang}(k, \lambda) \\ N &\sim \text{Poisson}(\lambda) \end{aligned}$$

The probability statements that are the same are:

$$\mathbb{P}(T_k > 1) = \mathbb{P}(N \leq k-1) = \mathbb{P}(N < K)$$

The interpretation of the l.h.s is if you are waiting for $k \overset{iid}{\sim} \text{Exp}(\lambda)$ events to occur / happen / realize in sequence then the probability of waiting for more than 1 second means that $k-1$ or less events happened before 1 second. Due to the equivalence, the Poisson r.v. N models the number of events that happen in less than 1 second. You can write this in CDF's as:

$$1 - F_{T_k}(1) = F_N(k-1)$$

(b) [easy] Write the assumptions and the main result of the Binomial-Negative Binomial relationships (an equivalence of two probability statements and then an equivalence using the CDF's of the Binomial and the Negative Binomial models).

(c) [easy] Assume $X_1, X_2, X_3, \ldots \overset{iid}{\sim} \text{Exp}(\lambda)$. Calculate $\mathbb{P}(X_1 + X_2 + X_3 + X_4 + X_5 < 1)$ using the two different ways (i.e. via the Poisson Process relationship).

We know by the convolution of exponentials that

$$X_1 + X_2 + X_3 + X_4 + X_5 = T_5 \sim \text{Erlang}(5,\, \lambda)$$

So one easy way to calculate this probability is:

$$\mathbb{P}(X_1 + X_2 + X_3 + X_4 + X_5 < 1) = \mathbb{P}(T_5 < 1) = F_{T_5}(1)$$

Using the $N \sim \text{Poisson}(\lambda)$ is slightly harder. Using our answer in (a), we can derive:

$$1 - F_{T_k}(1) = F_N(k-1) \quad \Rightarrow \quad 1 - F_N(k-1) = F_{T_k}(1)$$

which means that we can get our second answer as

$$\mathbb{P}(X_1 + X_2 + X_3 + X_4 + X_5 < 1) = 1 - F_N(4)$$

This is equivalent to $\mathbb{P}(N \geq 5) = \mathbb{P}(N = 5) + \mathbb{P}(N = 6) + \ldots$.

What is the intution here? If you are waiting for $k \overset{iid}{\sim} \text{Exp}(\lambda)$ events to occur / happen / realize in sequence then the probability of waiting for less than 1 second means that k or more events happened before 1 second.

(d) [easy] Let $N \sim \text{Poisson}(\lambda)$. Describe a way to se the realizations from the r.v.'s $X_1, X_2, X_3, \ldots \overset{iid}{\sim} \text{Exp}(\lambda)$ to create a realization n from the Poisson model.

(e) [difficult] Assume $X_1, X_2, X_3, \ldots \overset{iid}{\sim} \text{Exp}(\lambda)$. Calculate $\mathbb{P}(X_1 + X_2 + X_3 + X_4 + X_5 < m)$ where $m \in \mathbb{N}$ using two different ways (i.e. via the Poisson Process relationship).

Using our answer from (c), we can easily use the Erlang CDF again replacing the 1 with m:

$$\mathbb{P}(X_1 + X_2 + X_3 + X_4 + X_5 < m) = \mathbb{P}(T_5 < m) = F_{T_5}(m)$$

If $N \sim \text{Poisson}(\lambda)$ models the number of events in 1 second, then $N_1 + N_2 + \ldots + N_m$ model the number of events in m seconds. Since the exponentials are $\overset{iid}{\sim}$, the number occuring in each second are also $\overset{iid}{\sim}$ meaning $N_1, N_2, \ldots, N_m \overset{iid}{\sim} \text{Poisson}(\lambda)$ and we proved in a previous homework that a convolution of Poissons is Poisson, so $N_1 + N_2 + \ldots + N_m \sim \text{Poisson}(m\lambda)$.

So now we use our answer from (c) switching the single N model with our sum model:

$$\mathbb{P}(X_1 + X_2 + X_3 + X_4 + X_5 < m) = 1 - F_{N_1+N_2+\ldots+N_m}(4)$$

(f) [harder] Assume $X_1, X_2, X_3, \ldots \overset{iid}{\sim} \text{Geometric}(p)$. Calculate $\mathbb{P}(X_1 + X_2 + X_3 + X_4 + X_5 < 10)$ using two different ways (i.e. via the Binomial-Negative Binomial relationship).

Problem 3

These exercises will give you practice with transformations of discrete r.v.'s.

(a) [easy] Let $X \sim \text{Binomial}(n, p)$. Find the PMF of $Y = g(X) = \ln(X + 1)$.

(b) [harder] Let $X \sim \text{Binomial}(n, p)$. Find the PMF of $Y = g(X) = X^2$. Is $g(X)$ monotonic? Does that matter for this r.v.?

(c) [difficult] Let $X \sim \text{Binomial}(n, p)$ where n is an even number. Find the PMF of $Y = g(X) = \text{mod}(X, 2)$ where "mod" denotes modulus division of the first argument by the second argument.

(d) [difficult] [MA] Let $X \sim \text{NegBin}(k, p)$. Find the PMF of $Y = g(X) = \text{mod}(X, n)$ where $n \in \mathbb{N}$.

Problem 4

These exercises will give you practice with transformations of continuous r.v.'s and the quantile function.

(a) [harder] Let $X \sim U(0, 1)$. Find the PDF of $Y = g(X) = aX + c$. Make sure you're careful with the indicator function that specifies the support. There are two cases.

(b) [harder] Let $X \sim \text{Exp}(\lambda)$. Find the PDF of $Y = g(X) = \ln(X)$.

(c) [difficult] [MA] Let $X \sim \text{Exp}(\lambda)$. Find the PDF of $Y = g(X) = \sin(X)$.

(d) [harder] Let $X \sim \text{Exp}(1)$. Find the PDF of $Y = g(X) = -\ln\left(\frac{e^{-X}}{1-e^{-X}}\right)$. If this is a brand name r.v., mark it so and include its parameter values.

(e) [easy] Find the Quantile function of X where $X \sim \text{Logistic}(0, 1)$.

(f) [easy] Find the PDF of $Y = \sigma X + \mu \sim \text{Logistic}(\mu, \sigma)$ where $X \sim \text{Logistic}(0, 1)$.

(g) [difficult] Let $X \sim \text{Logistic}(0, 1)$. Find the PDF of $Y = g(X) = \frac{1}{1+e^{-X}}$. If this is a brand name r.v., mark it so and include its parameter values.

(h) [easy] Let $X \sim \text{Exp}(\lambda)$. Find the PDF of $Y = g(X) = ke^X$. Marked easy because it is in the notes. This will be a brand name r.v., so mark it so and include its parameter values.

(i) [easy] Rederive the $X \sim \text{Laplace}(0, 1)$ r.v. model by taking the difference of two standard exponential r.v.'s.

(j) [easy] Let $X \sim \text{Laplace}(0, 1)$. Prove that $\mathbb{E}[X] = 0$ without using the integral definition. There's a trick.

(k) [easy] Find the PDF of $Y = \sigma X + \mu \sim \text{Laplace}(\mu, \sigma)$ where $X \sim \text{Laplace}(0, 1)$.

(l) [harder] [MA] Find the Quantile function of X where $X \sim \text{Laplace}(0, 1)$.

(m) [harder] [MA] Prove that $X \sim \text{ParetoI}(1, log_4(5))$ models the "Pareto Principle".

(n) [harder] Is $X \sim \text{ParetoI}(1, log_4(5))$ a good model for land ownership amount for individuals? Why / why not?

(o) [difficult] [MA] Let $X \sim \text{ParetoI}(k, \lambda)$. Show that $Y = X \mid X > c$ where $c > k$ is also a ParetoI r.v. and find its parameter values.

(p) [E.C.] [MA] Prove or disprove that considering any ParetoI conditional on being larger than a certain value would also be a power rule where the top q proportion of the unit modeled is in the hands of the top $\bar{q} := 1 - q$ values of the unit.

This is the whole idea of a power law e.g. if the top 1% of the country owns 99% of the wealth then the power law also implies that of the top 1% of the top 1% of them owns 99% of that 99% and etc.

I couldn't seem to get it in one hour of trying so it is likely hard to show this. Or maybe it's incorrect. You get EC if you make a good effort.

Problem 5

We will now explore a couple of extreme distributions.

(a) [harder] [MA] Let $X \sim \text{Exp}(1)$ and $Y = -\ln(X) \sim \text{Gumbel}(0, 1)$. Find the PDF of this standard Gumbel distribution. Make sure you include the indicator function in the functional form to make it valid $\forall\, y \in \mathbb{R}$ (if necessary).

(b) [easy] Find the CDF of Y.

(c) [easy] Let $G = \beta Y + \mu \sim \text{Gumbel}(\mu, \beta)$. Find the PDF of G, the general Gumbel distribution.

(d) [easy] [MA] Show that for any r.v. X, if $Y = aX + b$, then $F_Y(y) = F_X\left(\frac{y-b}{a}\right)$.

(e) [easy] Using the answer in the previous question, find the CDF of $G \sim \text{Gumbel}(\mu, \beta)$.