Math 621 Fall 2017
Midterm Examination One

Solutions

Professor Adam Kapelner

October 3, 2017

Full Name ____

Code of Academic Integrity

Since the college is an academic community, its fundamental purpose is the pursuit of knowledge. Essential to the success of this educational mission is a commitment to the principles of academic integrity. Every member of the college community is responsible for upholding the highest standards of honesty at all times. Students, as members of the community, are also responsible for adhering to the principles and spirit of the following Code of Academic Integrity.

Activities that have the effect or intention of interfering with education, pursuit of knowledge, or fair evaluation of a student's performance are prohibited. Examples of such activities include but are not limited to the following definitions:

Cheating Using or attempting to use unauthorized assistance, material, or study aids in examinations or other academic work or preventing, or attempting to prevent, another from using authorized assistance, material, or study aids. Example: using an unauthorized cheat sheet in a quiz or exam, altering a graded exam and resubmitting it for a better grade, etc.

I acknowledge and agree to uphold this Code of Academic Integrity.

______________________________ signature   ______________ date

Instructions

This exam is seventy five minutes and closed-book. You are allowed one page (front and back) of a "cheat sheet." You may use a graphing calculator of your choice. Please read the questions carefully. If the question reads "compute," this means the solution will be a number otherwise you can leave the answer in choose, permutation, exponent, factorial or any other notation which could be resolved to a number with a computer. I advise you to skip problems marked "[Extra Credit]" until you have finished the other questions on the exam, then loop back and plug in all the holes. I also advise you to use pencil. The exam is 100 points total plus extra credit. Partial credit will be granted for incomplete answers on most of the questions. $\boxed{\text{Box}}$ in your final answers. Good luck!

Problem 1 Below are some theoretical exercises.

(a) [7 pt / 7 pts] Assume independent r.v.'s X and Y where $\mathbb{S}\text{upp}[X] = (0, \infty)$ and $\mathbb{S}\text{upp}[Y] = (0, \infty)$. Beginning from the general definition of convolution, prove that

$$f_{X+Y}(t) = \int_0^t f_X(x) f_Y(t-x)dx$$

where by convention, the notation f_X represents the PDF of r.v. X which *does not include* an indicator function.

$$f_{X+Y}(t) = \int_{\mathbb{R}} f_X(x)\, \mathbb{1}_{x \in Supp(X)}\, f_Y(t-x)\, \mathbb{1}_{t-x \in Supp(Y)}\, dx$$ def

$$= \int_{\mathbb{R}} f_X(x)\, \mathbb{1}_{x \in (0,\infty)}\, f_Y(t-x)\, \mathbb{1}_{t-x \in (0,\infty)}\, dx$$ by assumption

$$= \int_0^\infty f_X(x)\, f_Y(t-x)\, \mathbb{1}_{t-x \in (0,\infty)}\, dx$$ indicator function adjusts the bounds

$$= \int_0^t f_X(x)\, f_Y(t-x)\, dx$$ //

(b) [7 pt / 14 pts] If $X, Y \overset{iid}{\sim} \text{Exp}(\lambda)$, find $\mathbb{P}(X \geq Y)$.

$$\mathbb{P}(X \geq Y) = E[\mathbb{1}_{X \geq Y}] = \iint_{x \in Supp(X),\, y \in Supp(Y)} \mathbb{1}_{x \geq y} f_{X,Y}(x,y)\, dx\, dy = \int_{x=0}^{x=\infty}\int_{y=0}^{y=\infty} \lambda e^{-\lambda x} \lambda e^{-\lambda y} \mathbb{1}_{x \geq y}\, dx\, dy$$

$$= \int_{y=0}^{\infty} \lambda e^{-\lambda y} \int_{x=y}^{\infty} \lambda e^{-\lambda x}\, dx\, dy = \int_{y=0}^{\infty} \lambda e^{-\lambda y} \left[-e^{-\lambda x}\right]_y^\infty dy = \int_0^\infty \lambda e^{-\lambda y} e^{-\lambda y}\, dy$$

$$= \lambda \int_0^\infty e^{-2\lambda y}\, dy = \left[-\frac{1}{2} e^{-2\lambda y}\right]_0^\infty = \boxed{\frac{1}{2}}$$

(c) [10 pt / 24 pts] If $X, Y \overset{iid}{\sim} \text{Exp}(\lambda)$, find the conditional density of X given $X+Y$. There is no guarantee that the result will be the density of a brand name r.v., but if it is, denote it and find the parameter(s) as a function of λ.

$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$ let $T = X+Y \Rightarrow T \sim \text{Erlang}(2,\lambda)$ $$f_{X|T}(x,t) = \frac{f_{X,T}(x,t)}{f_T(t)} = \frac{f_{X,Y}(x,t-x)}{f_T(t)}$$

$$= \frac{\lambda e^{-\lambda x}\lambda e^{-\lambda(t-x)}}{\frac{\lambda^2 e^{-\lambda t} t^{2-1}}{(2-1)!}} = \frac{\lambda^2 e^{-\lambda t}}{\lambda^2 e^{-\lambda t} t} = \frac{1}{t} = \text{Unif}(0,t)$$

PDF of Erlang evaluated at t

Note: $\text{Supp}[X|T] = (0,t)$ because if $x+y=t \Rightarrow x \in (0,t)$

(d) [7 pt / 31 pts] Let $X \sim \text{Binomial}(n_1, p_1)$ independent of $Y \sim \text{Binomial}(n_2, p_2)$. Find the PMF of $X+Y$ *as best as you possibly can* (even if there is no closed form solution).

$$T = X+Y \sim \sum_{x \in \text{Supp}[X]} p_X(x)\, p_Y(t-x) = \sum_{x=0}^{n} \binom{n_1}{x} p_1^x (1-p_1)^{n_1-x} \binom{n_2}{t-x} p_2^{t-x}(1-p_2)^{n_2-t+x}$$

$$= p_2^t (1-p_1)^{n_1}(1-p_2)^{n_2-t} \sum_{x=0}^{n} \binom{n_1}{x}\binom{n_2}{t-x}\left(\frac{p_1(1-p_2)}{p_2(1-p_1)}\right)^x$$

(e) [7 pt / 38 pts] Let $X \sim$ Binomial (2000, 0.004) independent of $Y \sim$ Binomial (20000, 0.0004). Approximate $\mathbb{P}(X+Y=0)$. The correct answer uses one simple operation and no credit will be given for answers that require the use of non-trivial computing.

$X \sim Bin(2000, 0.004) \overset{d}{\approx} Poisson(2000 \cdot 0.004) = Poisson(8)$

$Y \sim Bin(20000, 0.0004) \overset{d}{\approx} Poisson(20000 \cdot 0.0004) = Poisson(8)$

$T = X+Y \overset{d}{\approx} Poisson(16)$ via the convolution done in class

$P(T=0) = \frac{e^{-16}\lambda^0}{0!} = \boxed{e^{-16}}$

(f) [5 pt / 43 pts] Let $X \sim$ NegBin (35, 0.37) independent of $Y \sim$ Geometric (0.37). Find the PMF of $X+Y$.

$T = X+Y \sim NegBin(36, 0.37) = \binom{x+35}{35} \cdot .63^x \cdot .37^{36}$

(g) [5 pt / 48 pts] Calculate $\gamma(5, 24.5) + \underbrace{\int_{24.5}^{\infty} t^4 e^{-t} dt}$ as a number $\in \mathbb{N}$ explicitly.

$\gamma(5, 24.5) + \Gamma(5, 24.5) = \Gamma(5) = 4! = 24$

Problem 2 Below are some questions about waiting times.

(a) [5 pt / 53 pts] The time until phone the next phone call is exponentially distributed with an average of half hour. If you have already waited half hour, find the probability you will wait more than another half hour.

$X \sim Exp(\lambda)$ $E(X) = \frac{1}{\lambda} = \frac{1}{2} \Rightarrow \lambda = 2$

irrelevant due to memorylessness property

$= P(X > \frac{1}{2} + \frac{1}{2} \mid X > \frac{1}{2})$

$= P(X > \frac{1}{2})$

$= 1 - F_X(\frac{1}{2})$

$= e^{-2\frac{1}{2}} = \boxed{\frac{1}{e}}$

(b) [5 pt / 58 pts] The time until phone the next phone call is exponentially distributed with an average of half hour. What is the probability you get two phone calls in one hour?

Via the Poisson process in one dimension, the # of events in one time unit

is $N \sim \text{Poisson}(\lambda)$. Here $\lambda = 2$

$$P(N=2) = \frac{e^{-2}\cancel{2}^{\cancel{2}}}{\cancel{2}!} = \boxed{\frac{2}{e^2}}$$

(c) [7 pt / 65 pts] The time until phone the next phone call is exponentially distributed with an average of half hour. What is the probability you get 10 phone calls before 5hr? You can answer using the CDF of a r.v. you define.

$T \sim \text{Erlang}(10, 2)$

$P(T \leq 5) = F_T(5)$

Problem 3 Below are some theoretical exercises about the vector-valued r.v.'s.

(a) [6 pt / 71 pts] Let $X_1, \ldots, X_n \overset{iid}{\sim}$ Binomial (n, p) and let $\boldsymbol{X}$ denote the vector of these r.v.'s. Find $\mathbb{V}\text{ar}[\boldsymbol{X}]$.

Since $\overset{iid}{\sim}$ all $Cov[X_i, X_j] = 0$ if $i \neq j$. The diagonal terms are $Var(X_i) = np(1-p)$

$\Rightarrow Var[\vec{X}] = np(1-p) I_n$ where I_n is the identity matrix of dimension n.

(b) [7 pt / 78 pts] Let $X_1, \ldots, X_n \overset{iid}{\sim}$ Binomial (n_0, p) and let $\boldsymbol{X}$ denote the vector of these r.v.'s. Is $\boldsymbol{X} \sim$ Multinomial $(n, \boldsymbol{p})$? If yes, find the values of its parameters as functions of n_0 and p. If no, explain why not.

Note: this should probably read Multinomial(m, pvec) not "n" since n was used above

No. $\sum_{i=1}^{n} X_i \in \{0, 1, \ldots, n n_0\}$ and <u>not fixed</u>.

(c) [7 pt / 85 pts] A person goes to the grocery store and buys n fruits. For each of his n selections, he picks rambutans with probability p otherwise he picks dragonfruits. If rambutans cost a and dragonfruits cost b, find his expected bill.

$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \text{Multi}\left(n, \begin{bmatrix} p \\ 1-p \end{bmatrix}\right)$ where X_1 denotes the # of rambutans and X_2 denotes the # of dragonfruits

let $\vec{c} = \begin{pmatrix} a \\ b \end{pmatrix}$ $E(B) = E\left[\vec{c}^T \vec{X}\right] = \vec{c}^T \vec{\mu} = \begin{pmatrix} a & b \end{pmatrix} n \begin{bmatrix} p \\ 1-p \end{bmatrix} = n \cdot \left(ap + b(1-p)\right)$

$B = \vec{c}^T \vec{X}$ denotes the bill

(d) [15 pt / 100 pts] Find the standard deviation of his bill and simplify as best as possible.

$SE[B] = \sqrt{Var(B)} = \sqrt{Var(\vec{c}^T \vec{X})} = \sqrt{\vec{c}^T Var(\vec{X}) \vec{c}} = \sqrt{np(1-p) \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}} = |a-b| \sqrt{np(1-p)}$

$Var(\vec{X}) = \begin{bmatrix} np(1-p) & -np(1-p) \\ -np(1-p) & np(1-p) \end{bmatrix} = np(1-p) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a & b \end{pmatrix} \begin{bmatrix} a-b \\ -a+b \end{bmatrix} = a^2 - ba - ab + b^2 = a^2 - 2ab + b^2 = (a-b)^2$