Math 368 / 621 Fall 2020
Midterm Examination Two

Professor Adam Kapelner

Wednesday, November 4, 2020

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.

By taking this exam, you acknowledge and agree to uphold this Code of Academic Integrity.

Instructions

This exam is 75 minutes (variable time per question) and closed-book. You are allowed **one** page (front and back) of a "cheat sheet", blank scrap paper and a graphing calculator. Please read the questions carefully. No food is allowed, only drinks.

Problem 1 [4min] (and 4min will have elapsed) Let $Y_1, ..., Y_n \overset{iid}{\sim} \text{Poisson}\,(\lambda_Y)$ be the discrete rv that models the number of runs the Yankees score in n individual games and $M_1, ..., M_n \overset{iid}{\sim} \text{Poisson}\,(\lambda_M)$ be the discrete rv that models the number of runs the Mets score in n individual games. Let the difference in scores for any game i be $D_i := Y_i - M_i$.

- [5 pt / 5 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

(a) Any D_i follows a Skellam distribution without additional assumption(s).

If you marked (a) false, for the rest of this problem assume the assumption(s) needed so that D_i is Skellam.

(b) $D_i \sim \text{Skellam}(Y, M)$

(c) $D_1, ..., D_n$ are $\overset{iid}{\sim}$

(d) $D_1 + \ldots + D_n$ follows a Skellam distribution

(e) $M_i - M_j$ where $i \neq j$ follows a Skellam distribution.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 2 [7min] (and 11min will have elapsed) Let $X \sim U(\{1, 4, 9, 16, 25, 36\})$ and $Y = \sqrt{X}$ and $Z = \mathrm{mod}(X, 2)$.

- [8 pt / 13 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

  (a) X is a discrete rv

  (b) Y is a continuous rv

  (c) $p_Y(y) = p_X(y^2)$

  (d) $p_Y^{old}(y) = 1/6$

  (e) $\exists h$ such that $p_Z(z) = p_X(h(z))$

  (f) Z is a Bernoulli rv

  (g) $p_Z^{old}(z) = 1/2$

  (h) The rv's Y and Z are independent.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 3 [11min] (and 22min will have elapsed) Let $X \sim \text{Gamma}(\alpha, \beta) := f_X(x)$ and $Y = 1/X \sim f_Y(y)$.

- [15 pt / 28 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

(a) $F_X(x) = \int_{\mathbb{R}} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx$

(b) $F_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^x u^{\alpha-1} e^{-\beta u} du$

(c) $F_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \gamma(x, \alpha)$

(d) $F_X(x) = P(x, \alpha)$

(e) $F_X(x) = P(\alpha, \beta x)$

(f) $F_Y(y) = 1 - F_x(1/y)$

(g) $F_Y(y) = F_x(1/y)$

(h) When doing the kernel decompositions $f_X(x) = c_X k_X(x)$ and $f_Y(y) = c_Y k_Y(y)$ we see that $c_X = c_Y$.

(i) When doing the kernel decompositions $f_X(x) = c_X k_X(x)$ and $f_Y(y) = c_Y k_Y(y)$ we see that $k_X(x) = k_Y(y)$.

(j) XY is distributed as a standard uniform.

(k) The support of Y is all real numbers.

(l) $f_Y^{old}(y) = \frac{\Gamma(\alpha)}{\beta^\alpha} (1/y)^{\alpha+1} e^{-\beta/y}$.

(m) $f_Y^{old}(y) = \frac{\Gamma(\alpha)}{\beta^\alpha} (1/y)^{\alpha-1} e^{-\beta/y}$.

(n) $f_Y^{old}(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} (1/y)^{\alpha+1} e^{-\beta/y}$.

(o) $f_Y^{old}(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} (1/y)^{\alpha-1} e^{-\beta/y}$.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 4 [10min] (and 32min will have elapsed) Let $X_1, X_2, \ldots \overset{iid}{\sim} \text{Exp}(\lambda)$ and $N \sim \text{Poisson}(\lambda)$.

- [12 pt / 40 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

(a) $\int_{\mathbb{R}} p_N(t)dt = 1$.

(b) $\int_{\mathbb{R}} f_{X_1}(t)dt = 1$.

(c) $\int_{\mathbb{R}} f_{X_1+X_2+\ldots+X_k}(t)dt = 1$ for all $k \in \mathbb{N}$.

(d) X_1 is distributed as an Erlang rv.

(e) X_1 is distributed as a Gamma rv.

(f) $X_1 + X_2 + \ldots$ is distributed as an Erlang rv.

(g) $X_1, X_2, \ldots$ and N are related by the Poisson Process without further assumptions.

If you marked (g) false, for the rest of this problem assume whatever is necessary so that $X_1, X_2, \ldots$ and N are related by the Poisson Process.

(h) $\mathbb{P}(X_1 + X_2 + \ldots + X_k < t) = P(k, \lambda t)$

(i) $\mathbb{P}(N < t) = P(k, \lambda t)$

(j) $\mathbb{P}(X_1 + X_2 + \ldots + X_k \geq 1) = \mathbb{P}(N < k)$.

(k) $\mathbb{P}(X_1 + X_2 + \ldots + X_k \leq 1) = \mathbb{P}(N \geq k)$.

(l) $\mathbb{P}(X_1 + X_2 + \ldots + X_k \leq 1) + \mathbb{P}(N < k) = 1$.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 5 [6min] (and 38min will have elapsed) Let $X \sim \text{Logistic}(\mu, \sigma)$ and let $Z = \frac{X-\mu}{\sigma}$.

- [8 pt / 48 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

  (a) $f_{X,Z}(x, z) = f_X(x) f_Z(z)$

  (b) Z has support of all the real numbers.

  (c) $\mathbb{V}\text{ar}\left[Z\right] = 1$

  (d) $\mathbb{E}\left[X\right] = \sigma$

  (e) X could be an *error distribution*

  (f) $F_X(1/2) = \mu$

  (g) The 50%ile of Z is $\ln(1)$

  (h) The 90%ile of Z is $\ln(9)$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 6 [4min] (and 42min will have elapsed) Let $X_1, X_2, ..., X_n \overset{iid}{\sim} \text{Laplace}(0, 1)$.

- [5 pt / 53 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

  (a) $f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \frac{1}{2^n} e^{-|x_1 + \ldots + x_n|}$

  (b) $Q[X_1, q] = 1 - Q[X_1, 1 - q]$

  (c) X_1 could be an *error distribution*

  (d) $X_1 + X_2 + \ldots + X_n$ could be an *error distribution*

  (e) $X_1 + X_2 + \ldots + X_n$ is Laplace distributed

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 7 [8min] (and 50min will have elapsed) Let $X \sim \text{Weibull}(k, \lambda)$ and $Y = \ln(X)$.

- [6 pt / 59 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

  (a) X could be a memoryless rv.

  (b) The kernel of $f_X(x)$ is $k\lambda^k$.

  (c) if $k = 1.00001$ then $\mathbb{P}(X \geq x + c \mid X \geq c) < \mathbb{P}(X \geq x)$ for all $c > 0$.

  (d) Since X is a waiting time / survival distribution, Y is also a waiting time / survival distribution.

  (e) $f_Y(y)^{old} = e^{k\lambda(\lambda y)^{k-1}e^{-(\lambda y)^k}}$

  (f) $f_Y(y)^{old} = k\lambda^k e^{ky} e^{-\lambda^k e^{ky}}$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 8 [11min] (and 61min will have elapsed) Let $X_1, X_2, \ldots, X_{17} \overset{iid}{\sim} U(0,1)$.

- [12 pt / 71 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

(a) $X_1, X_2, \ldots, X_{17}$ are called "order statistics".

(b) The third largest realization is denoted $x_{(3)}$.

(c) The third largest realization is denoted $X_{(3)}$.

(d) $\mathbb{S}\text{upp}\left[X_{(k)}\right] = [0, k)$.

(e) The probability that the third largest is less than 0.3 can be computed via the function $\frac{1}{B(3,15)}x^2(1-x)^{14}$ by plugging in $x = 0.3$.

(f) In the density of the rv modeling the maximum of $X_1, X_2, \ldots, X_{17}$, most of the mass is located near zero.

(g) In the density of the rv modeling the maximum of $X_1, X_2, \ldots, X_{17}$, most of the mass is located near 0.5.

(h) In the density of the rv modeling $X_{(9)}$, most of the mass is located near zero.

(i) In the density of the rv modeling $X_{(9)}$, most of the mass is located near 0.5.

(j) If $Y_1 = aX_1, Y_2 = aX_2, \ldots, Y_{17} = aX_{17}$ where a is a positive constant, the density of the minimum of $Y_1, Y_2, \ldots, Y_{17}$ is $\frac{17}{a}\left(\frac{y}{a}\right)^{16}\mathbb{1}_{y \in [0,a]}$.

(k) If $a < 1$, $b > 1$, $c > 2$, then $B(a, b, c) < B(b, c)$.

(l) If $a = 0$, $b > 1$, $c > 2$, then $I_a(b, c) > B(b, c)$.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 9 [9min] (and 70min will have elapsed) Let $X_1, X_2 \overset{iid}{\sim} \text{Exp}(\lambda)$ and $Y_1 = \ln(X_1/X_2)$. And let $\boldsymbol{g}$ denote a vector valued function from two input x's to two output y's, $\mathbb{R}^2 \to \mathbb{R}^2$ and $\boldsymbol{h}$ denote the multivariate inverse function for $\boldsymbol{g}$ which is thus also a vector valued function from two input y's to two output x's, $\mathbb{R}^2 \to \mathbb{R}^2$. Let $g_1(x_1, x_2) = \ln(x_1/x_2)$ and $g_2(x_1, x_2) = x_2$.

- [8 pt / 79 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

  (a) $h_1(y_1, y_2) = \ln(y_1/y_2)$ and $h_2(y_1, y_2) = y_2$ fits the above definition of $\boldsymbol{h}$.

  (b) $f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(y_2 e^{y_1}, y_2) y_2 e^{y_1}$

  (c) $f_{Y_1,Y_2}(y_1, y_2) = f_X(y_2 e^{y_1}) f_X(y_2) y_2 e^{y_1}$

  (d) $f_{Y_1,Y_2}(y_1, y_2) = \lambda^2 y_2 e^{-\lambda y_2 (1+e^{y_1})}$

  (e) There is one nuisance dimension in the jdf of $\boldsymbol{Y}$ when we wish to find the density of Y_1.

  (f) $f_{Y_1}(y) = \lambda^2 y e^{-\lambda y} \int_{\mathbb{R}} e^{-\lambda y e^u} du$

  (g) $f_{Y_1}(y) = \lambda^2 \int_{\mathbb{R}} u e^{-\lambda u (1+e^y)} du$

  (h) Y_1 will be BetaPrime distributed.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 10 [5min] (and 75min will have elapsed) Let $X \mid N = n \sim \text{Binomial}\,(n,\, 1/2) := \binom{n}{x}\frac{1}{2^n}$ and let n be a realization from $N \sim \text{ExtNegBin}(\alpha, p)$.

- [7 pt / 86 pts] Record the letter(s) of all the following that are **true**. At least one will be true.

  (a) The rv $X \mid N$ is discrete.

  (b) The rv X (unconditional on N) is discrete.

  (c) The rv X is a compound distribution.

  (d) The rv X (unconditional on N) will have two parameters.

  (e) The rv X (unconditional on N) will have three parameters.

  (f) The old-style PDF or PMF for X (unconditional on N) can be expressed $\int_{\mathbb{R}} \binom{n}{x}\frac{1}{2^n}\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)n!}(1-p)^n p^\alpha dn$.

  (g) The old-style PDF or PMF for X (unconditional on N) can be expressed $\sum_{n=0}^{\infty} \binom{n}{x}\frac{1}{2^n}\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)n!}(1-p)^n p^\alpha$.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.