STAT 134 - Instructor: Adam Lucas

Midterm 2

SOLUTIONS

Friday, November 20, 2020

Print your name: ________________________________

SID Number: ________________________________

Exam Information and Instructions:

- You will have 48 hours to take this exam. Open book/notes but no internet resources outside of the Stat134.org website. You are allowed to use a caluculator.
- We will be using Gradescope to grade this exam. Write any work you want graded on the front of each page, in the space below each question. Additionally, write your SID number in the top right corner on every page.
- Provide calculations and reasoning in every answer.
- Unless stated otherwise, you may leave answers as unsimplified numerical and algebraic expressions, and in terms of the Normal c.d.f. Φ. Finite sums are fine, but simplify any infinite sums.

I certify that all materials in the enclosed exam are my own original work and I have not violated the UC Berkeley honor code.

Sign your name: ________________________________

GOOD LUCK!

SID: ______________________

1. (1 pt) Circle all choices that violate the rules of this test or the UC Berkeley Honor Code?

   (a) post or read on Chegg or an online forum

   (b) use a calculator

   (c) leaving your answer unsimplified

   (d) use your notes, or textbook

   (e) communicate with non-staff about the test until we communicate with you that everyone has taken the test.

   Solution: a,e

2. New MGF

   Sankar and Zhiyi are unsatisfied with the moment-generating function, and invent a new function called the S-function. The S-function of a random variable X is defined as:

   $$S_X(t) = M_{\log(X)}(t)$$

   Where $M_{\log(X)}(t)$ is the moment-generating function of $\log(X)$.

   Now let X, Y be independent and identically distributed as per the following density:

   $$f(z) = \frac{1}{z^2}, 1 \leq z < \infty$$

   (a) (3 pts) Show that $S_{\mathbf{X}}(t) = \frac{1}{1-t}$, $t < 1$.

   Hint: You do not need to do a change of variables for this.

   Answer.

   $$S_X(t) = M_{\log(X)}(t) = E[e^{\log(X)t}] = E[X^t] = \int_1^\infty x^{t-2}dx = -\frac{1}{t-1} = \boxed{\frac{1}{1-t},\ t < 1}$$

   (b) (3 pts) Find $S_{\mathbf{XY}}(t)$, which is the S-function of the product of X and Y, and make sure to define where it is finite.

   Answer.

   $$S_{XY}(t) = E[e^{\log(XY)t}] = E[e^{(\log(X)+\log(Y))t}]$$

   $$= E[e^{\log(X)t}e^{\log(Y)t}] = E[e^{\log(X)t}]E[e^{\log(Y)t}] = S_X(t)S_Y(t) = \boxed{\left(\frac{1}{1-t}\right)^2,\ t < 1}$$

3. Sum of Independent Normals and Distance, change of variables

   Suppose that at each timestep t, a drunken ant moves randomly in the x and y directions such that the movements are independent of one another across direction and time. Let $X_t, Y_t \sim \mathcal{N}(0, \sigma^2)$.

(a) (2 points) Find the distribution of $\sum_{i=1}^{n} X_t$.

Answer. Since any linear combination of independent normals is normal, $\boxed{\sum_{t=1}^{n} X_t \sim \mathcal{N}\left(0, n\sigma^2\right)}$.

(b) (4 points) Find the probability density function for,

$$\sqrt{\left(\sum_{t=1}^{n} X_t\right)^2 + \left(\sum_{t=1}^{n} Y_t\right)^2},$$

the distance of the ant from its starting point at time step n.

Answer. Note that:

$$\sqrt{\left(\sum_{t=1}^{n} X_t\right)^2 + \left(\sum_{t=1}^{n} Y_t\right)^2} = \sqrt{n\sigma^2}\sqrt{\left(\frac{\sum_{t=1}^{n} X_t}{\sqrt{n\sigma^2}}\right)^2 + \left(\frac{\sum_{t=1}^{n} Y_t}{\sqrt{n\sigma^2}}\right)^2}$$

Since the terms inside the square root are the sums of squares of independent standard normals, the inner square root has the Rayleigh distribution. Then, by change of scale (i.e. change of variable rule) the density of the entire quantity is:

$$\frac{1}{\sqrt{n\sigma^2}}\left(\frac{r}{\sqrt{n\sigma^2}}\right)\exp\left\{-\frac{1}{2}\left(\frac{r}{\sqrt{n\sigma^2}}\right)^2\right\}, \left(\frac{r}{\sqrt{n\sigma^2}}\right) > 0 = \boxed{\frac{r}{n\sigma^2}\exp\left\{-\frac{1}{2}\frac{r^2}{n\sigma^2}\right\}, r > 0}$$

(c) What does this density look like as the number of time steps n gets larger, and what does this say about where the ant may end up in the distant future? (A sound logical argument suffices. No calculations needed.)

Answer. Note that as n gets large for a fixed σ^2, the density becomes more and more "dispersed", so that more probability is concentrated in successively longer tails. This suggests that as n gets large, we have more uncertainty about the ant's final distance/position with respect to its starting point.

4. Competing Exponentials

After learning of the new P/NP grading options, you think about the chance of being able to (separately) call your two closest friends to share the news, before your phone battery dies.

Assume that:

- the calls last i.i.d. $\text{Exp}(\lambda)$ amounts of time;
- the phone battery has an $\text{Exp}(\mu)$ amount of life remaining when you begin making calls; and
- the call durations and the battery life are independent.

You quickly calculate the probability that you complete both calls before your phone dies, as follows.

Let X be the battery life and Y_1 and Y_2 the durations of the first two calls.

$$\begin{aligned}\mathbb{P}(X \geq Y_1 + Y_2) &= \mathbb{P}(X \geq Y_1 + Y_2 | X \geq Y_1) \cdot \mathbb{P}(X \geq Y_1) \\ &= \mathbb{P}(X \geq Y_2) \cdot \mathbb{P}(X \geq Y_1) \\ &= \mathbb{P}(X \geq Y_1)^2 \\ &= \left(\frac{\lambda}{\mu + \lambda}\right)^2.\end{aligned}$$

Is your derivation correct? If yes, justify each step. If not say what is wrong.

Answer.

If we denote by X the battery life and by Y_1 and Y_2 the durations of the first two calls, then the probability in question is

$$\begin{aligned}\mathbb{P}(X \geq Y_1 + Y_2) &= \mathbb{P}(X \geq Y_1 + Y_2 | X \geq Y_1) \cdot \mathbb{P}(X \geq Y_1) \\ \text{(by memorylessness)} \quad &= \mathbb{P}(X \geq Y_2) \cdot \mathbb{P}(X \geq Y_1) \\ \text{(by } Y_1 \stackrel{d}{=} Y_2\text{)} \quad &= \mathbb{P}(X \geq Y_1)^2 \\ \text{(by competing exponentials)} \quad &= \left(\frac{\lambda}{\mu + \lambda}\right)^2.\end{aligned}$$

In words, for two calls to end before the battery dies, we need the first call to end before the battery dies. This is an exponential race, which is handled by competing exponentials. Having guaranteed this, memorylessness implies that the probability with which the battery lasts an additional amount of time sufficient to complete another call is the same as if we tried to complete that call at time zero. Because the calls have identical distributions, this is like doing the first call over again, which is again handled by competing exponentials.

Common mistakes.

1) Concerning the first equality, many students neglected to observe that it is not just the total probability rule. That would be

$$\mathbb{P}(X \geq Y_1 + Y_2) = \mathbb{P}(X \geq Y_1 + Y_2 | X \geq Y_1)\mathbb{P}(X \geq Y_1) + \mathbb{P}(X \geq Y_1 + Y_2 | X < Y_1)\mathbb{P}(X < Y_1),$$

as the events $\{X \geq Y_1\}$ and $\{X < Y_1\}$ are disjoint and exhaustive (i.e., one of them occurs). One must also observe that the second term is zero, as X cannot be at least $Y_1 + Y_2$ (which both take nonnegative values) unless $X \geq Y_1$.

2) Concerning the second equality, several students suggested that memorylessness does not apply because $Y_1 + Y_2$ is not an exponential random variable. Memorylessness, however, is being applied to X. Suppose the values of Y_1 and Y_2 were known to be t and s, then this is simply

$$\mathbb{P}(X \geq t + s | X \geq t) = \mathbb{P}(X \geq s),$$

which is the memorylessness property applied to X.

3) Concerning the third equality, many students erroneously stated it held due to "symmetry."

4) Concerning the fourth equality, many students tried to use the exponential tail of $\mathbb{P}(X \geq x) = e^{-\mu x}$. This is not necessarily misguided, as one can condition Y_1 to be x, in which case $e^{-\mu x}$ is the right probability, but the conditioning must then be removed by integrating against the density of Y_1. This calculation leads to the competing exponentials formula.

5. Convolution Suppose there are two i.i.d random variables $U_1, U_2 \sim \text{Uniform}(0,1)$. Let $X = \max\{U_1, U_2\}$, $Y = \min\{U_1, U_2\}$.

   (a) (3 points) Derive the joint density of X and Y and show your work.
   Answer. To find the joint density of $Y = U_1$ and $X = U_2$ we can find

$$P(X \in dx, Y \in dy) \approx f(x,y)dxdy,$$

   by throwing two darts at the unit interval with $0 < y < x < 1$ as done in class. This will lead you immediately to $P(X \in dx, Y \in dy) \approx \binom{2}{1,1} dxdy$. It follows that $f(x,y) = 2$.
   Alternatively,

$$\begin{aligned} \mathbb{P}(X \leq x, Y \leq y) &= \mathbb{P}(X \leq x) - \mathbb{P}(X \leq x, Y > y) \\ &= x^2 - (x-y)^2 \end{aligned}$$

   Therefore, the joint density $f_{X,Y}(x,y) = \frac{d^2}{dxdy}\mathbb{P}(X \leq x, Y \leq y) = 2$, where $x \geq y$.

   (b) (3 points) Prove via the convolution formula that the density of $Z = X - Y$ is Beta distribution, and provide the parameters.
   The density of Z via convolution formula is

$$\begin{aligned} f_Z(z) &= \int_0^{1-z} f_{X,Y}(y+z, y)dy \\ &= 2(1-z), \quad z \in [0,1] \end{aligned}$$

   which is $Beta(1,2)$.

6. (Waiting times),

   (a) A family is getting ready for their trip to Yosemite. Each person is in their room, packing their bags. For each person, the time it takes them to pack their bag is exponentially distributed and independent of the time it takes any other person. On average, it takes each parent 1 hour and each child 2 hours to get ready. In a family with 2 parents and 4 children, what is the probability that it takes the family more than 2 hours to get ready?

Solution (non–international): We are asked to compute the probability that $M = \max(X_1, X_2, X_3, X_4, X_5, X_6)$ where X_1, X_2 are $\exp(1)$ and X_3, X_4, X_5, X_6 are $\exp(1/2)$, is at most 2. We have that

$$\mathbb{P}(M \le x) = \mathbb{P}(X_1, \dots, X_6 \le x) = \mathbb{P}(X_1 \le x) \cdots \mathbb{P}(X_6 \le x) = (1 - e^{-x})^2 \left(1 - e^{-\frac{x}{2}}\right)^4.$$

So the probability that it takes them more than 2 hours is

$$\mathbb{P}(M > 2) = 1 - \mathbb{P}(M \le 2) = 1 - (1 - e^{-2})^2 \left(1 - e^{-1}\right)^4.$$

(b) After setting up the tents the family wants to rent bicycles. The rental place does only have one bicycle when they get there. Assume that the bikes being returned follow a Poisson process with rate $\frac{1}{10}$ (measured in minutes) and that nobody else is waiting for bikes. Give an expression for the probability that the family will have to wait more than 30 minutes before they can start their biking tour. (No need to simplify!)

Solution (non–international): If B_t is the number of bikes returned by time t this question is asking for the probability that B_{30} is smaller or equal to 4. B_{30} is Poisson $30\frac{1}{10} = 3$. Thus

$$\mathbb{P}(B_{30} \le 4) = \sum_{k=0}^{4} e^{-3} \frac{3^k}{k!}.$$

Alternatively, with T_i denoting the arrival time of the ith bike, we have that T_i is $Gamma(i, \frac{1}{10})$, so $\mathbb{P}(B_{30} \le 4) = \mathbb{P}(T_5 > 30) = \int_{30}^{\infty} \frac{x^4 e^{-\frac{x}{10}}}{4! 10^5}$

Solution (international): If B_t is the number of bikes returned by time t this question is asking for the probability that B_{40} is smaller or equal to 4. B_{40} is Poisson $40\frac{1}{20} = 2$. Thus

$$\mathbb{P}(B_{40} \le 4) = \sum_{k=0}^{4} e^{-2} \frac{2^k}{k!}.$$

Alternatively, with T_i denoting the arrival time of the ith bike, we have that T_i is $Gamma(i, \frac{1}{20})$, so $\mathbb{P}(B_{40} \le 4) = \mathbb{P}(T_5 > 40) = \int_{40}^{\infty} \frac{x^4 e^{-\frac{x}{20}}}{4! 20^5}$

SID: ____________________________

Scratch Paper

SID: ____________________

Discrete

name and range	$P(k) = P(X = k)$ for $k \in$ range	mean	variance
uniform on $\{a, a+1, \ldots, b\}$	$\frac{1}{b-a+1}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
Bernoulli (p) on $\{0, 1\}$	$P(1) = p$; $P(0) = 1 - p$	p	$p(1-p)$
binomial (n, p) on $\{0, 1, \ldots, n\}$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson (μ) on $\{0, 1, 2, \ldots\}$	$\frac{e^{-\mu}\mu^k}{k!}$	μ	μ
hypergeometric (n, N, G) on $\{0, \ldots, n\}$	$\frac{\binom{G}{k}\binom{N-G}{n-k}}{\binom{N}{n}}$	$\frac{nG}{N}$	$n\left(\frac{G}{N}\right)\left(\frac{N-G}{N}\right)\left(\frac{N-n}{N-1}\right)$
geometric (p) on $\{1, 2, 3 \ldots\}$	$(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
geometric (p) on $\{0, 1, 2 \ldots\}$	$(1-p)^k p$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
negative binomial (r, p) on $\{0, 1, 2, \ldots\}$	$\binom{k+r-1}{r-1} p^r (1-p)^k$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$

SID: ______________________

Continuous

† undefined.

name	range	density $f(x)$ for $x \in$ range	c.d.f. $F(x)$ for $x \in$ range	Mean	Variance
uniform (a, b)	(a, b)	$\frac{1}{b-a}$	$\frac{x-a}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
normal $(0, 1)$	$(-\infty, \infty)$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$	$\Phi(x)$	0	1
normal (μ, σ^2)	$(-\infty, \infty)$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}$	$\Phi\left(\frac{x-\mu}{\sigma}\right)$	μ	σ^2
exponential (λ) = gamma $(1, \lambda)$	$(0, \infty)$	$\lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$
gamma (r, λ)	$(0, \infty)$	$\Gamma(r)^{-1} \lambda^r x^{r-1} e^{-\lambda x}$	$1 - e^{-\lambda x} \sum_{k=0}^{r-1} \frac{(\lambda x)^k}{k!}$ for integer r	r/λ	r/λ^2
chi-square (n) =gamma $\left(\frac{n}{2}, \frac{1}{2}\right)$	$(0, \infty)$	$\Gamma(\frac{n}{2})^{-1} (\frac{1}{2})^{\frac{n}{2}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$	as above for $\lambda = \frac{1}{2}$, $r = \frac{n}{2}$ if n is even	n	$2n$
Rayleigh	$(0, \infty)$	$x e^{-\frac{1}{2}x^2}$	$1 - e^{-\frac{1}{2}x^2}$	$\sqrt{\frac{\pi}{2}}$	$\frac{4-\pi}{2}$
beta (r, s)	$(0, 1)$	$\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} x^{r-1} (1-x)^{s-1}$	see Exercise 4.6.5 for integer r and s	$\frac{r}{r+s}$	$\frac{rs}{(r+s)^2(r+s+1)}$
Cauchy	$(-\infty, \infty)$	$\frac{1}{\pi(1+x^2)}$	$\frac{1}{2} + \frac{1}{\pi}\arctan(x)$	†	†
arcsine =beta $(1/2, 1/2)$	$(0, 1)$	$\frac{1}{\pi\sqrt{x(1-x)}}$	$\frac{2}{\pi}\arcsin(\sqrt{x})$	$\frac{1}{2}$	$\frac{1}{8}$